© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

SECTION I

Foundation Competencies

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

Chapter 2

Working with Numbers

LEARNING OBJECTIVES

1. To be able to calculate and use descriptive statistics.
2. To be able to compare different types of data using statistical inference and hypothesis testing.
3. To be able to present data effectively and efficiently in visual form.

REAL WORLD SCENARIO

James Walden is the CEO of Port City Hospital, a medium-sized hospital located in a moderately populated seacoast region in New England. It services a community of approximately 230,000 people, but has four competitor hospitals within a 30-mile radius. Over the past 10 years, civic development leaders have been transforming a closed air force base into a trade center and have been slowly attracting corporate entities to the area. In the past year, two major employers have relocated their corporate offices and/or large manufacturing centers to the trade center, adding an estimated 30,000 new persons to the area. In addition, another large insurance company has doubled its workforce in the same town as Port City Hospital, creating an additional 4000 jobs, at least two thirds of which they anticipate filling from outside the area. Within the hospital, Mr. Walden has been hearing increased concerns from physician and nursing leadership about workload increases. In looking at utilization data, Mr. Walden is curious if the hospital is realizing a growth pattern that is both different than previous years and different than other hospitals in the area. Knowing this might facilitate his decision to expand certain services or the hospital itself.

Numbers are a form of language. For managers, they facilitate the way we communicate about the functioning of organizations. They allow us to count, describe, compare, and predict. Mathematics and statistics—the calculation and manipulation of numbers—are the primary tools that facilitate these numerical “conversations.”

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

Numbers and analysis also provide justification for action (or sometimes inaction). They are static and unbiased, and when used properly and transparently, they allow those interpreting their analysis to make reasoned and informed decisions. Health services managers utilize numbers extensively in their work, because most of what is done in any organization must be quantified. As the adage goes; if you can't measure it, you can't manage it.

This chapter reviews the basic forms of mathematical and statistical analysis used in health services administration, including various descriptive and inferential statistics. The chapter concludes with basic guidelines and examples of how to best present data in tabular and graphical form.

LEARNING OBJECTIVE 1: CALCULATING AND USING DESCRIPTIVE STATISTICS

Statistics is a term that usually invokes dread and discomfort in students and practitioners alike. This is probably because statistics holds a close relationship to mathematics, although the two are distinct. Statistics uses mathematical relationships between data to allow managers to make decisions about both the data itself and about the likelihood that sampled data represent a broader and more generalized trend or population. Statistics, however, need not be difficult. For managers, some simple categorizing of techniques can help focus statistical analysis in ways that are easily understood and applied.

Managerial statistics have three primary functions. The first function is to describe certain data elements, such as the number of births over a time period or the expenses incurred for a service unit. The second function is to compare two points of data, such as births from 1 year to the next or error rates between care sites. The third is to predict data, such as visit volume in future months. This chapter examines the first two, saving prediction for Chapters 5, 6, and 7, given its specialized nature. First, however, a discussion about the nature of data is in order. Data are quite simply numbers within a context. Green, although a very nice color, is only an adjective by itself. If, however, we record the eye color of a room of 20 people, and then code those colors with numbers—i.e., 1 for blue, 2 for brown, 3 for green—we have transformed the colors into points of data. Similarly, if we were to then count the number of people with green eyes, we have performed a statistical function—the calculation of a descriptive statistic.

A number of different types of analyses can be performed on data. When the time comes to conduct these analyses, students often face "analysis paralysis." Imagine, for example, that you find yourself in a new position, and you have been asked by your boss, Mr. Walden, to take a look at some utilization trends using data from the organization's data warehouse. You pull up the corresponding data files, and are faced with more than 30 variables (columns in a spreadsheet) and

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

tens of thousands of records (rows in a spreadsheet representing individual patient visits). In the middle is a sea of numbers of various sorts that continue on as you scroll and scroll down the screen. In some organizations, there may be millions of data records in thousands of tables, which is intimidating to be sure. However, a data file with 10,000 rows and one with 20 are not all that different. Each can be described in similar ways. What is important is the data itself. Understanding what the numbers represent (the context) and how they were created will lead you down certain analytic paths and not others, allowing you to put some statistical methods aside for some data.

Measuring Data

Data come in only four varieties. Students of introductory statistics will no doubt recall the terms *nominal*, *ordinal*, *interval*, and *ratio*. All refer to measurement of data variables. Variables are simply data that can take on different values, depending on what is being measured. In the earlier example, the color of 20 people's eyes was recorded, thus creating the variable "eye color." In this instance it is a variable that is measured nominally. Nominal refers to data that exist in non-overlapping categories. They have no ranking and are mutually exclusive; for example, eye color, insurance type, gender, and ethnicity. Ordinal variables are slightly different in that they are still measured categorically, but the categories have a ranking. An example of this would be satisfaction scales, in which somewhat satisfied might be followed by very satisfied, etc. These are common in health surveys. The final two types are often taken together as interval/ratio variables. These are often termed *continuously measured variables*; examples include time and money. The difference here is that they are actually still categories, but the distance between categories is equal. Think of a time scale as derived in seconds—one second, two seconds, etc. We could derive smaller increments if we so wished, creating fractions of seconds as is often done in Olympic time trials and racing. The increments ultimately do not matter, however. What does matter is that the distance between them is equal. This allows mathematical calculations on these forms of data. The difference between interval and ratio data has to do with the presence of a meaningful zero when measuring a ratio variable, a distinction not important for this discussion.

At this point the insightful student might realize that examining measurement provides two distinct types of variables—those that have equal distances between measurement points and those that do not. Often, these distinctions are recognized by labeling nominal and ordinal data as categorical, and interval/ratio data as continuous. We too will follow this convention. Ordinal data present a unique measurement form. It is important to understand the type of data you are working with because each is analyzed differently.

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

Descriptive Statistics with One Variable (Univariate)

First, examine descriptive statistics in relation to categorical data. Table 2-1 provides data on 14 patients, recording their insurance type.

Insurance type is a categorical variable. The categories are not ranked, nor is there any relationship among them. Patients usually claim a type of primary insurance (or lack thereof) upon visit. To describe the data, we are limited to only a handful of techniques. The first is to simply count. Here we can count total patients or patients by the type of insurance they have. The second, which requires a bit of mathematics to be conducted first, is to create percentages for the number of persons falling into each category. A percentage is simply the number of persons in a category divided by the total number of persons, multiplied by 100. Not multiplying by 100 is also correct, although this provides a decimal fraction and not a percentage. For example, we may wish to know how many people reported having United as their insurer. One way to summarize this would be to count, which amount to three individuals in Table 2-1. To calculate a percentage, we would divide that 3 by 14, which gives us 0.21, or 21% of patients. Percentages and fractions provide slightly more information than do counts. Inherent in them is the context of the whole. If we tell you three patients had United insurance, you may still wonder if that is a lot, not many, or a modest amount; but if we say 21% of patients had United, you now have some sense of the entire group of patients, although we have not provided the total. Here, providing the total in addition to the percentage would provide both the count and the total, creating a more complete picture of the data being described. Listing counts and percentages of categorical data is also called creating frequencies from the data. We could graph the data at this point and obtain a visual representation of how frequently patients used various types of insurance as we have done in Figure 2-1. From a descriptive standpoint, this is the limit of analyzing a singular categorical variable.

Table 2-1 Insurance Type by Patient

1	United	8	BC/BS
2	Medicare	9	Medicaid
3	Medicaid	10	Uninsured
4	Medicare	11	Medicare
5	BC/BS	12	Uninsured
6	United	13	United
7	BC/BS	14	MBCA

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

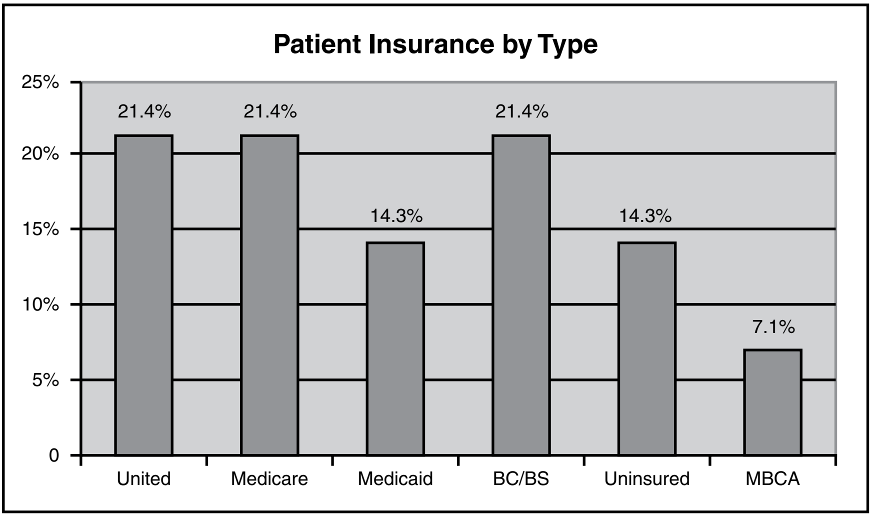

Figure 2-1 Patient Insurance by Type

The second type of data you may wish to describe are those measured as interval/ratio variables. These are also commonly referred to as continuous data. Again, these are actually categorical data as well, but the categories are of equal size. Examples include variables such as time, money, height, and weight. The equal distances between categories are what allow for mathematical analysis of these data. So, for example, adding one dollar to two dollars adds the same amount as adding one dollar to ten dollars. This allows us to calculate a number of descriptive measures that examine the centrality of the data and its spread, which are both useful for our purposes. We first examine measures of central tendency.

Measures of Central Tendency

As described, data that are collected across a number of observations vary from observation to observation; thus, the term *variable*. Plotting these data reveals both the spread and the clustering of individual observations. An example is given in Figure 2-2.

From these data, we can see that the number of chart pulls appear to largely be centered between 10 and 30 per day, with a few days of higher volume, and one with lower volume. What would be helpful for analytic purposes would be to have a set of summary statistics to describe the data. The statistic that is the mathematical center of a data set is the average, or *mean*. It is the foundation for many other statistical concepts as well. To calculate the *mean*, simply add up all

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

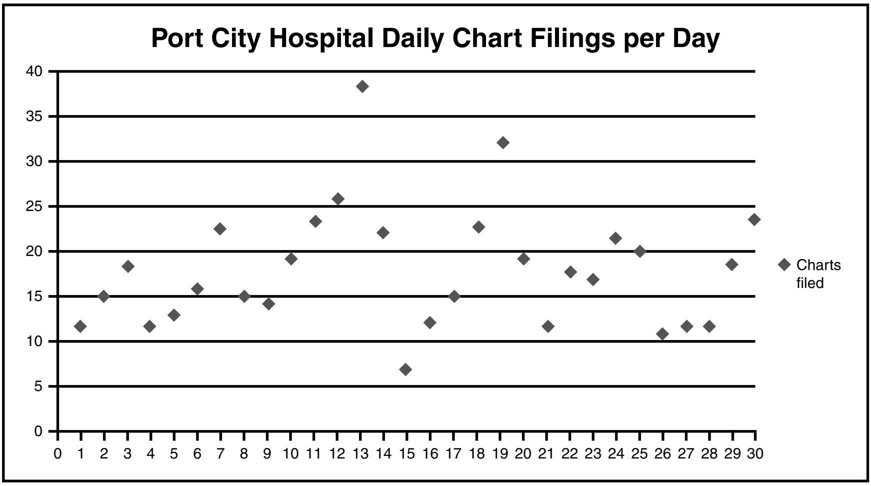

Figure 2-2 Port City Hospital Daily Chart Filings per Day

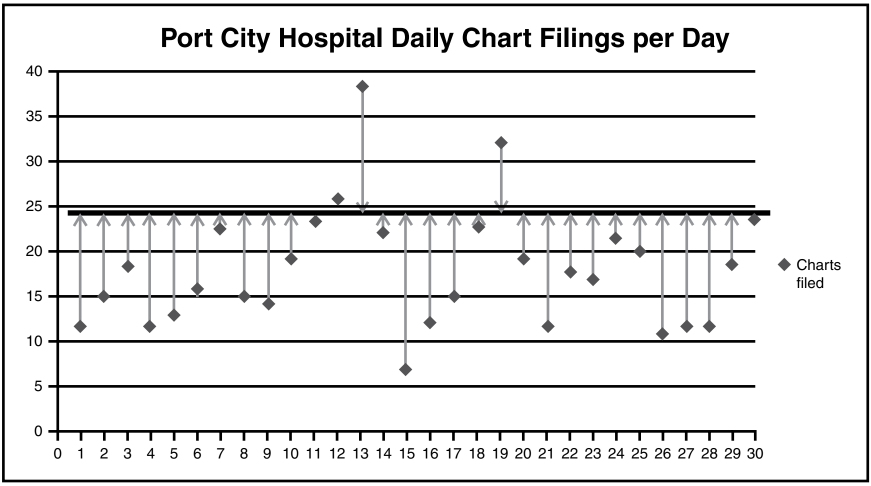

Figure 2-2A Port City Hospital Daily Chart Filings per Day Part 2

the values and divide by *n*, which is the number of observations. We can also find the *median*, which is the center of the distribution of data when all the observations are arranged from lowest to highest. The *mode* is the more frequently reported data value. Given our data from Figure 2-2, Table 2-2 reports these measures of central tendency.

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

Table 2-2 Port City Hospital Daily Chart Filings per Day

Day	*Charts Filed*	*Day*	*Charts Filed*
1	12	16	12
2	15	17	15
3	18	18	23
4	12	19	32
5	13	20	19
6	16	21	12
7	22	22	18
8	15	23	17
9	14	24	21
10	19	25	20
11	23	26	11
12	26	27	12
13	38	28	12
14	22	29	18
15	7	30	23
Mean	**18**		
Median	**18**		
Mode	**12**		

In this instance, the median and mean are the same value, 18. The mode is 12. Had the mean been higher than the median, it would indicate that there were some high values of the data that were pulling the mean upward. Examine the following range of income values: $13,000, $25,000, $33,000, $42,000, $56,000. The mean of these data is $33,800. The median is $33,000. If, however, we replace the value of $56,000 with $120,000, notice what happens. The median is still $33,000, yet the mean increases to $46,600. This is because the median is not dependent on all other values in the distribution. It is what we call a *robust measure*, or one that is resistant to outlying values. The mean is not robust, as we demonstrated. When examining data distributions, it is sometimes helpful to look at both the mean and the median. Doing so can quickly tell you something about the presence of outlying values and the spread of the data.

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

Measures of Spread

Although the mean, median, and mode tell us something about the middle or centrality of the data, we may also be interested in how varied and spread out the data are. This is both helpful to understand the range of data values, and also to examine the possibility of outlier values that might be affecting our measures of central tendency. Examine again the data in Table 2-2 and Figure 2-2. We know that the mean of these data is 18 charts filed on average. We also know that there are many days when the number of charts filed exceeds 18 per day and also falls short of 18 per day. The maximum and minimum values tell us this, and are important measures for summarizing our data. Here they are 48 and 7, respectively. Their difference, or 41 $(48-7)$ is what is known as the *range*. Examining the range in addition to other measures of central tendency allows a clearer picture of the data distribution (even without a graph!).

There is one final measure of spread that should be considered. If we were to draw a line at the mean in Figure 2-2 we would see that about half of the data points were clustered above and half below (and although this is always true of the median, it need not always be so for the mean) (see Figure 2.2 A). Here we see that some points lie closer to the mean than others, whereas some lie on the mean. Thus, each point of observation lies some distance from the mean, whether positive or negative. What would be interesting is to know how far from the mean are the data *on average*. The final summary measure of spread does this, which is the standard deviation. Simply put, the standard deviation is the average distance of a given data point to its mean. In the chart filing example, we are asking on average how far do the data points diverge from their mean, which is 18? To do this we could start by measuring the distance of each point to the mean, and then simply dividing by *n* to get the average. But wait. Because the mean is the mathematical average of all the points, the distances when summed will total zero. Dividing zero by anything is a mathematical impossibility. So, to counter this problem, the negative distances need to be eliminated by squaring all of the distances. This eliminates our zero total problem, but also converts all our original distances into *squared* distances, so that when we add them up and divide by *n*, we have the average total squared distance, also known as the *variance*. In this case, this creates a measure interpreted as the number of charts pulled squared. This creates an interpretive problem in that we no longer have the same units with which we started. To return to our original units requires that we eliminate the squared term by taking the square root, thus providing the standard deviation.

Working with Samples

The preceding calculation provides the standard deviation for a set of data. If those data constitute a complete set of observations, and generalization to some larger

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

population is not being made; for example, from a sample of chart filings to estimate all chart filings, then the standard deviation should be calculated in this way. However, if we are using a sample value, which is known, to say something about a population value, which usually cannot be known, we must make an adjustment. Samples are inherently more variable than populations. We are simply more likely to get data points further away from the "true" population mean in a sample than were we to continue to collect more and more data. Because of this variability, when calculating our standard deviation, we divide by $(n - 1)$.

When dealing with sample data, we also need to be careful when interpreting the mean of the data. Table 2-2 is a sample of data for 1 month of chart filings. If we are only interested in that month, we can treat the data as a population. However, if we want to treat this 1 month as representative of all months, some adjustment is required. The mean of the data in Table 2-2 was 18 chart filings. If we were to resample these data over another time period, what is the likelihood that 18 would again be the mean? If we designate 18 as a sample value representative of the "truth," we are in fact saying that it *is* and *always will be* 18. This is quite unlikely in this instance. However, it is often not possible for us to know the "truth" for all present and future data. Instead we can create an interval that we can say with some level of confidence contains the "true" population mean. The formula for constructing a confidence interval at the 95% level of confidence, our default for most analyses, is:

$$\text{Mean} \pm 1.96 \times \text{standard error}$$

where the standard error = standard deviation / square root (n).

The value of 1.96 is the value that cuts off the upper and lower 2.5% of the standard normal distribution (discussed briefly later in this chapter) and the use of the standard error rather than the standard deviation is to adjust for the fact that we are using a sample (with greater variability) to represent a population. The reporting of confidence intervals should be included with any mean that has been derived from sample data.

LEARNING OBJECTIVE 2: TO BE ABLE TO COMPARE DIFFERENT TYPES OF DATA USING STATISTICAL INFERENCE AND HYPOTHESIS TESTING

Bivariate Analysis

The second primary function of managerial statistics is to be able to compare two or more variables within a set of data. This can mean comparing a variable measured at two points in time, such as the number of births from one year to the next, or in two locations, such as comparing births between hospitals. It can also

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

mean comparing two different types of data, such as the number of emergency department (ED) visits over a time period with the number of lab tests performed during that same period. Each type of analysis again requires knowing what types of data you are comparing. Like descriptive statistics, there are certain types of analyses that you will perform and others we can set aside depending on how the data are measured. Before doing so, however, we first need to review the need for hypothesis development and testing.

Hypothesis Testing

Students may recall from an introductory statistics course, that comparisons of variables are best tested using hypotheses. These are simply statements of association that are first stated and then, using analysis, either supported or refuted. The reason for doing this is that most data are simply representations of phenomena that exist in real life. The problem is that it is usually unrealistic or impossible to measure all phenomena completely. Think about measuring population. We may want to know the actual number of people in the United States, but our ability to measure this is limited. Realistically, we cannot find and count everyone in the United States without missing some people, and the number of people changes daily because of births and deaths, so that by the time we were done measuring, the "real" answer would have already changed. Yet, we also know that for a given point in time there is a "real" measurement; we just are unable to observe it. We *can*, however, estimate within some level of certainty, whether or not the measurement we observe, or the data comparison we make, is likely to be representative of what is "real" at that point in time.

This is why we create hypotheses and then use statistical tests to either support or refute them. Two primary types of hypotheses are used in statistical analysis. The first is the null hypothesis, which is always the hypothesis of no association or difference. The second is the alternative hypothesis, which is most often the converse of the null, but that can be directional, such as two data elements having a positive or negative association. Both are examined here briefly.

Managers in healthcare settings are often assessing data for comparative purposes, and often they use samples of data taken at a point in time. The question managers should be concerned with is not only if there is an observable difference or association in the data, but with what level of confidence can you believe it to be true and not because of chance. Otherwise stated, if the manager collected another sample of data, could the association or difference reverse itself or would the data be reflecting the same pattern? These are the foundational questions behind hypothesis testing. For example, consider a manager who has collected data on the number of safety protocol violations within two units of the hospital. In summarizing these violations, she finds the mean number of violations of unit A to

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

be 21 over a 1-year period, and 27 in unit B over the same period. In real terms, unit B does report more violations than unit A. The question is whether this trend is a "real" trend, or simply a result of chance. Thus, the question we ask is how likely are we to record or see a difference as big as we have when the difference is actually zero in reality.

Often, the stating of hypotheses is unduly confusing. In fact, it is quite simple. The null hypothesis always states that there is no difference or association between the two things being observed (i.e., data variables). For example, we could hypothesize that the mean number of violations in site A is actually no different than the mean number of violations at site B "in reality" if we were to continue to measure over time. The alternative is that there is a real difference between the two things being observed. But here we have an option. We can say that we think, for example, that the mean number of violations at site B is simply different than site A, whether that is higher or lower. When direction doesn't matter, we are conducting a two-tailed hypothesis test. Our second alternative is to say that one will be higher than the other or lower than the other. In this case we might say the alternative hypothesis is that the mean number of violations at site B is higher than the mean number of violations at site A. In this case we are conducting a one-tailed hypothesis test. The difference occurs primarily with respect to interpretation of the tests and is explored later in the chapter. The stated null hypothesis, abbreviated H_0, or that of no difference, for this example would be:

H_0: There is no difference in the mean number of safety violations between site A and site B over the 1-year period.

The stated alternative hypothesis, abbreviated H_a, is the converse of this and, assuming a two-tailed test, would be:

H_a: There is a difference in the mean number of safety violations between site A and site B over the 1-year period.

The analysis we perform will allow us to either reject or fail to reject our null hypothesis. The rule here is that we *never* accept a hypothesis. Why? Because we can never be 100% certain what the relationship between two things is "in reality" at a given point in time, for reasons stated earlier in this section. Instead, we use hypothesis testing and statistics to make probabilistic inference into the relationship between two sets of measured data or observations. Interpreting our hypotheses now requires the use of statistics, and also a brief introduction to theoretical probability distributions, otherwise thought of as *why we can be certain we are at least partially certain.*

Probability Distributions

If someone were to ask you what the probability of flipping a normal coin and having it come up heads, you would no doubt say that it is a 50/50 chance, or 50%

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

of the time. Yet you would also likely agree that it is quite possible that you could flip a coin and heads would come up three times in a row. How can this be? Two reasons. One is that each coin flip is not dependent on the previous one. There are two sides of the coin, so you only have two possible outcomes. Each time you flip, they are equally likely to come up (if the coin is balanced and not a trick coin). The second is that we know if you continue to flip over and over again, the number of heads and the number of tails will start to equal out. In statistical language, we would say the probability of heads grows closer to 0.5 as your *n* (number of flips) increases. Suffice it to say that flipping a coin has a known probability. Could we observe 37 heads in a row? Sure, but it is highly unlikely.

Most phenomena in the world have a distribution of measurement, whether height, weight, income, hair length, etc. Consider height. There are a range of heights of individuals throughout the world. Some are quite tall, and others are not. If, for example, we see someone who is 8 feet tall, we might think that it is unusual, but not impossible. (Seeing is believing.) But how do we test this statistically?

Here we offer a non-statistical explanation of a statistical occurrence. As we stated before, at a point in time, all phenomena are theoretically measurable. If we are examining a data element that is continuous in nature, such as height, then at a point in time there is also a “true” mean of the observed data—right now there is a “true” mean height of all people in the world. Similarly, if we were to measure all persons, there would also be a “true” standard deviation around that mean. Some measurements will be close to the mean and others further away. We would expect that observations that were further from the mean would be less likely to occur, as with our 8-foot friend. From statistics we know how likely certain data will be to occur in relation to its mean by measuring how far those observations are from the mean in units of standard deviation. The reason for this is that many types of data are distributed normally, or in a fashion in which there is a mean and a symmetrical distribution of values on either side in the shape of a bell curve (Figure 2-3A). For example, *if* we were to know that the “true” mean of heights in the United States for men is 68 inches, and the “true” standard deviation is 2 inches, someone who is 8 feet tall (96 inches) would be 14 standard deviations above the mean or $(96-68)/2$. And because all data can be examined by how far in standard deviations they are from the mean, we can construct a theoretical normal distribution in which the mean is zero and the area under the curve represents units of standard deviation, called z-values. Why is the mean zero? If the distances under the curve are measures of standard deviation, then how many standard deviations away from the mean is the mean? Zero. Not only does this allow us to assess the probability of occurrence for certain data, it allows us to compare any type of data because the units are the same (standard deviations) (see Figure 2-3, A,B).

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

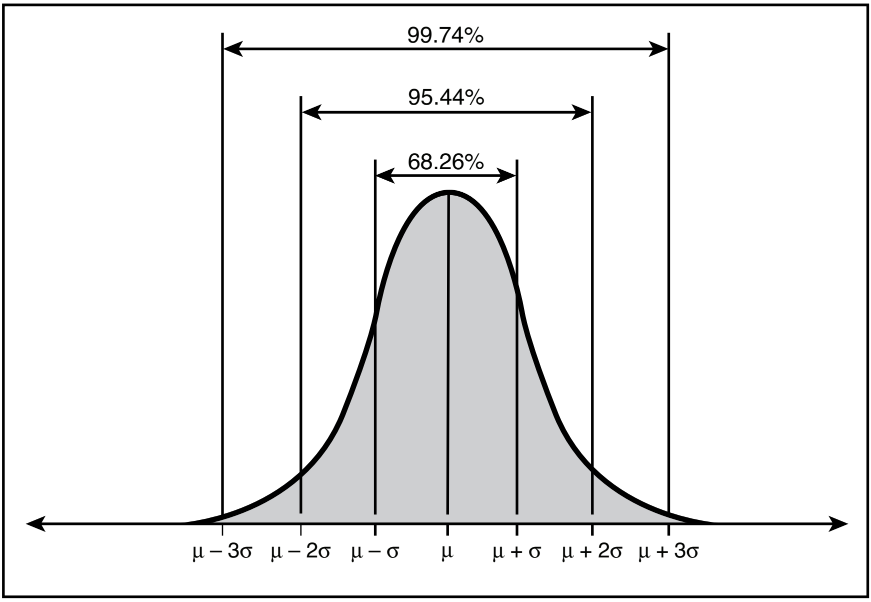

Figure 2-3A Normal Distribution Showing Mean, Standard Deviations, and Percentage of Observations Falling Within the Standard Deviations

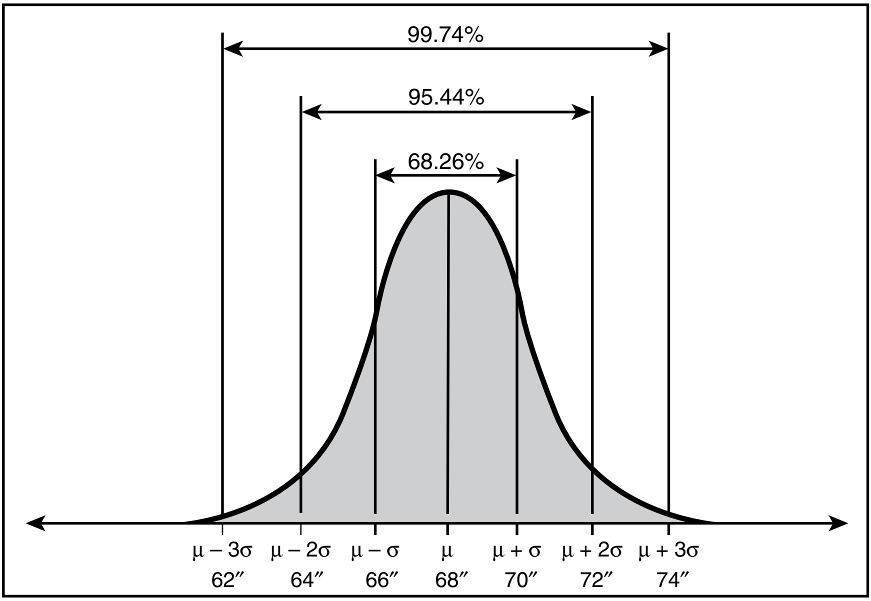

Figure 2-3B Distribution of Height (M = 68″ and SD = 2″)

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

Further, we know that the areas under standard normal distributions have known probabilities. The 68-95-99.7 rule states that approximately 68% of observations fall within one standard deviation of its mean, approximately 95% of observations fall within two standard deviations of the mean, and approximately 99.7% of observations fall within three standard deviations of the mean. In addition, each *z*-value under the curve has a known probability of occurrence. There are also other distributions that do not quite follow the symmetrical shape of a *z*-distribution (standard normal distribution), but nonetheless have known probabilities under them, such as *t*-distributions and *f*-distributions to name only two. For our purposes, it is not important to know what these distributions look like, but that they have known probabilities, and so any observations we may make can be tested to see what the likelihood of its occurrence is. All we need to know is what test to run to for what type of data, and then how to interpret the results we get from our computer analysis. We next examine this for a number of data comparisons.

Comparing Continuous Data

There are different analytic techniques for comparing a continuous data variable measured at different points in time or across locations, and two different continuous variables to one another. First, let us examine comparing two different continuously measured variables, lab tests and ED visits.

Correlation

In this instance, we look to statistics to provide a measure of association between two differently measured phenomena. Because they are measured in increments of equal distance, respectively, we can assess how unit changes in one variable are correlated to unit changes in the other. This becomes an algebraic relationship, in which if we label one variable x and the other y, we can express y as being some function of x. Examine Table 2-3, which measures the number of both ED visits and lab tests for a sample period during the month of September. If they were perfectly correlated, these variables would have a one-to-one relationship. In this example, a perfectly positive correlation would mean each additional ED visit would result in the same additional number of lab tests. Similarly, if there was a perfectly negative relationship, for every ED visit, lab tests would consistently decrease by a set amount. To do this we calculate the linear correlation coefficient (r), which will indicate the associative, but not causal relationship between the number of ED visits and the number of lab tests performed per day. The correlation coefficient will also indicate the strength of the linear association between the two variables.

Statistics indicates that a correlation coefficient (r) of $+1.00$ indicates a perfectly positive correlation (an increase in X is always associated with a parallel

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

Table 2-3 Observational Data for ED Visits and Lab Tests

Date	*Ed Visits*	*Lab Tests*
15-Sep	5	12
16-Sep	8	12
17-Sep	9	24
18-Sep	12	36
19-Sep	14	48
20-Sep	2	0
21-Sep	4	0
22-Sep	8	12
23-Sep	7	12
24-Sep	12	36
25-Sep	14	48
26-Sep	6	12
27-Sep	18	60
28-Sep	12	36

Correlation Coefficient
$r = .977$

Critical value of $r = .532$
with $n = 14$

increase in Y) and that a negative correlation coefficient of -1.00 indicates a perfectly negative correlation. By definition, correlation coefficients can only range from $+1.00$ to -1.00. For the data in Table 2-3, the correlation coefficient is 0.977, which indicates a strong positive correlation between ED visits and lab tests. Although this seems to be indication of a powerful correlation, the association cannot yet be said to be statistically significant or not one because of random chance.

Here we have collected a sample of data based on 14 days of observation. The question we must ask is whether the observed phenomenon could be owing simply to chance rather than some real association. Stating our hypotheses is helpful in doing this. Here, our null and alternative hypotheses are:

H_o: *There is no relationship between the number of ED visits and the number of lab tests.*

H_a: *There is a relationship between the number of ED visits and the number of lab tests.*

To address our hypotheses, we must now determine a critical value of r to assess the likelihood of the relationship being because of chance. Although some computer programs give the actual probability, or likelihood of the relationship with a

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

Table 2-4 Critical Values of the Correlation Coefficient *r* for Various Sample Sizes *n*

n	*r*	*n*	*r*
5	0.878	18	0.468
6	0.811	19	0.456
7	0.754	20	0.444
8	0.707	22	0.423
9	0.666	24	0.404
10	0.632	26	0.388
11	0.602	28	0.374
12	0.576	30	0.361
13	0.553	40	0.312
14	0.532	50	0.279
15	0.514	60	0.254
16	0.497	80	0.220
17	0.482	100	0.196

p-value, others, such as Excel, do not. Here we present a table of critical *r*-values, shown in Table 2-4, which gives the critical values of *r* at various sample sizes (n) at the alpha of 0.05. Given our example, we would use the critical value of *r* for $n = 14$, which is 0.532. If *r*-calculated > *r*-critical, we can be 95% confident that the association is not because of random chance. Note that analysis programs such as Excel will compute the correlation coefficient *r*, but do not provide the critical value of *r*, or the probability associated with *r*. Here, *r*-calculated (0.979) is greater than *r*-critical (0.532), so we can say that there is a statistically significant positive correlation between ED visits and lab tests. For each additional ED visit (1 unit), we would expect lab test volume to increase by 0.977 units.

T-Tests

A second common analysis is to examine a continuously measured variable at two points in time, or in two locations. For example, say we wish to compare the number of births as Port City Hospital with other U.S. hospitals of similar size. To do so we collect data over 12 months, shown in Table 2-5.

Examining the data we see that for all months, Port City Hospital performs more births than the average hospital of similar size in the United States and that the

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

mean number of births over the period was 37 at Port City and 29 at other hospitals. Our question is whether the data we see here for 1 year represent the "true" relationship between Port City and other similar size hospitals. Because this is only a sample of data from 1 year, we must use statistics to assess this. First, however, we should state our hypotheses.

H_o: *There is no difference between the mean number of births at Port City Hospital and other U.S. hospitals of similar size.*

H_a: *There is a difference between the mean number of births at Port City Hospital and other U.S. hospitals of similar size.*

To test the difference between two means requires the use of a *t*-test. *T*-tests are used to compare means between groups. These groups can be paired, as would be the case with a group who is measured on some variable, for example, blood pressure, undergoes some intervention, for example, an exercise routine, and remeasured. The groups can also be different, as is the case with our comparison of mean births at Port City Hospital with other hospitals. What cannot be compared are means for different variables, such as the mean average length of stay compared with the mean number of births. The means must be measured in similar units for comparison with a *t*-test.

Table 2-5 Comparative Monthly Births

	Port City Hospital	*U.S. for Similar Size Hospitals*
January	24	22
February	25	21
March	33	26
April	35	27
May	37	31
June	38	25
July	41	36
August	35	27
September	45	39
October	39	35
November	42	34
December	50	23
Mean	***37***	***29***

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

Most analytic software including Excel can calculate a number of *t*-tests. What is important to note is that the different types of *t*-tests (paired, assuming equal variances, and assuming unequal variances) revolve, as the names suggest, around variation of the data. For our purposes, we will assume that variances are unequal in cases other than paired data. In practice, this difference in variances would be analyzed with an *f*-test, and some programs will provide output for both equal and unequal variances assumed. Because Excel does not do this, assume unequal variances. Rarely will the interpretation differ between the two, but it can. The *t*-test output for our data is shown in Table 2-6.

In Table 2-6 we are given a number of analytic outputs. The first is the mean of the data for both Port City and the United States. We are also given the variance and the number of observations. The hypothesized mean difference is simply the null hypothesis restated. Examining the lower half of the table we are given the *t*-statistic, the probability of *t* for both one-sided and two-tailed tests, and the critical value of *t* that cuts off the upper *or* lower 2.5% of the distribution (one-tailed) and the value of *t* that cuts off the upper *and* lower 2.5% of the distribution (two-tailed). Thus, values of the *t*-statistic that lie beyond the critical value are statistically significant (different) at the 95% level of confidence. The *p*-value gives the exact probability that our means are truly different and that the observed difference is not because of random chance. Here, we would use a two-tailed *p*-value because our hypothesis was not directional. That is, our null stated that the mean number of births was different, but not in which direction (greater than or less than). Doing so would require a one-tailed test, because we would only be interested in values

Table 2-6 Excel Output *t*-Test: Two-Sample Assuming Unequal Variances

	Port City	*U.S.*
Mean	37	28.8
Variance	56	36.0
Observations	12	12
Hypothesized Mean Difference	0	
df	21	
t Stat	2.9499	
P(T≤*t*) one-tail	0.0038	
t Critical one-tail	1.7207	
P(T≤*t*) two-tail	0.0076	
t Critical two-tail	2.0796	

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

at one end of the distribution. Here it doesn't matter, so we conduct and interpret the two-tailed test. Interpreting the *t*-statistic we see that:

$$\text{T Stat } (2.95) > t \text{ Critical two tail } (2.07) \textit{ or } p(\text{T} \leq t)(0.0007) < 0.05$$

In either case, we would reject the null hypothesis.

Otherwise stated, our hypotheses ask what is the likelihood of seeing a difference in observed means as large as the one we did (8.2, or 37−28.8) if in fact the real difference were zero (the null hypothesis). Examining our *t* Stat relative to the critical values tells us the likelihood is less than 5% of the time. Examining our *p*-value tells us the exact likelihood, or less than 0.7% of the time (0.007). So, if we continue to collect samples of data, the means would likely only be the same in 0.7% of samples.

Comparing Categorical Data

Finally, we may often be interested in analysis of two variables that are measured categorically through either rates or proportions. Examples of these types of data include: What type of insurance does the patient have; what is the gender; and were they satisfied with their visit? Summarizing these involves creating counts and percentages. However, often we wish to compare how two groups of categories compare with one another. We do this using the chi-square statistic (χ^2), which compares the observed differences in proportions with what would be expected if proportions were equal. For example, if we were to examine the satisfied/unsatisfied percentages of 40 men and 40 women on a satisfaction questionnaire, we would expect that if they were equal, 20 would say satisfied and 20 would not in each gender category (or 50% for each). When we observe actual data, however, we often see different results. The basic null and alternative hypotheses hold true here. The question we are asking is what is the chance of seeing a difference of the magnitude observed in the collected data if in fact there is no true difference (all proportions are equal) in the population. The chi-square statistic and its associated probability allow us to test these hypotheses. The simplest form of chi-square analysis is of two variables using a 2 × 2 contingency table, shown in Table 2-7.

Table 2-7 2 × 2 Contingency Table

	Group 1	*Group 2*	*Total*
variable 1	*a*	*b*	*a* + *b*
variable 2	*c*	*d*	*c* + *d*
Total	*a* + *c*	*b* + *d*	*a* + *b* + *c* + *d*

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

Table 2-8 Patient Satisfaction Comparison Using Chi Square

	East Campus	*West Campus*	*Total*
Satisfied	36	17	53
Not satisfied	30	35	65
Total	66	52	118

Examine the data in Table 2-8 that depict satisfaction responses from a survey at two campuses of a clinic group. The null hypothesis would state that there is, in truth, no actual difference between satisfied and unsatisfied respondents by campus location. The alternative would be that the difference observed is real.

To calculate the chi-square, we use formula 2-1:

$$X^2 = \Sigma\left(\frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}\right), \text{ where the expected count is}$$

$$\frac{(\text{Row Total} \times \text{Column Total})}{n}$$

This formula yields the results in Table 2-9 and the chi-square statistic of 5.69. Constructing the expected values will be further helpful when using Excel to calculate the chi-square statistic.

This generates a chi-square statistic that must then be examined relative to the distribution of chi-squares for the given degree of freedom. Degrees of freedom are calculated by taking the number of rows minus one multiplied by the number of columns minus one. In this example we have one degree of freedom, or $(2 - 1) \times (2 - 1) = 1$. We now examine Appendix Table 1 to assess where our chi-square statistic falls relative to a given alpha level. We can see that for one

Table 2-9 Chi-Square Calculations for Patient Satisfaction Data

	Observed	*Expected*	*O − E*	*(O − E)²*	*(O − E)²/E*
	36	29.6	6.40	40.96	1.38
	17	23.4	−6.40	40.96	1.75
	30	36.4	−6.40	40.96	1.13
	35	28.6	6.40	40.96	1.43
Total	118	118	0.00	163.84	**5.69**

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

degree of freedom, a chi-square statistic of 5.61 falls beyond the alpha of 0.02. This means that we would be likely to see a difference in proportions of this magnitude when in fact they were equal less than 2% of the time. If our cutoff for significance was set at 5%, we would reject the null hypothesis in this instance.

In practice, most computer programs provide the chi-square statistic and corresponding significance when examining two or more categorical variables. When examining any categorical variable with more than two response categories, such as satisfaction levels or agreement scales, the chi-square statistic has a slightly different interpretive meaning. The null hypothesis remains the same in this case. However, we now observe differences not just between two categories (one and two), but between multiple categories (two and three, one and three, three and four, two and four, etc.). The chi-square can only tell us if the differences overall between the categories is significantly different than what we would expect, but does not test the differences between individual categories. Such analysis is beyond the scope of this book.

SUMMARY

Here we have presented a toolkit of basic statistical techniques to help guide the health services manager with basic quantitative analysis. This was not meant to be an exhaustive statistical review, but an applied, user-friendly introduction to statistics commonly used in making many healthcare related decisions. The first step in any analysis is to determine whether one is examining one variable or data point, or comparing more than one variable. When examining one variable at a time, we use descriptive statistics, and depending on whether the variable is continuous or categorical, different analyses are used. Table 2-10 provides a summary of these techniques for both describing and comparing data types.

Performing the right type of analysis on the data at hand can often be confusing for many, which is why we have attempted to segment analysis by the type of data being examined. These analyses will be referred to throughout the remainder of the text.

Table 2-10 Comparative Statistics Summary Table

<table>
<tr><td rowspan="2">Descriptive</td><td colspan="2">Continuous</td><td colspan="2">Categorical</td></tr>
<tr><td colspan="2">mean, median, mode, standard deviation, range, variance</td><td colspan="2">counts, percents, rates and proportions</td></tr>
<tr><td rowspan="2">Comparisons</td><td colspan="2">Continuous</td><td colspan="2">Categorical</td></tr>
<tr><td>Same variable</td><td>Different variable</td><td>Same variable</td><td>Different variable</td></tr>
<tr><td>Continuous</td><td>t-test</td><td>correlation</td><td>-</td><td>t-test</td></tr>
<tr><td>Categorical</td><td>-</td><td>t-test</td><td>chi-square</td><td>chi-square</td></tr>
</table>

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

LEARNING OBJECTIVE 3: TO BE ABLE TO PRESENT DATA EFFECTIVELY AND EFFICIENTLY IN VISUAL FORM

An important part of relaying messages gleaned from data is being able to effectively create visual representations of your data and your analyses. If we think about our presentation of data similarly to the way our data were analyzed; that is, descriptions of one point of data or variable at a time, and comparisons of data, we can examine a few simple tools and foundations for effective presentations. This section is by no means an effort to exhaust the functional abilities of graphical computer programs such as Microsoft Excel, but to provide a few basic tenets. Here we present uses of tables, bar/column graphs, pie charts, line graphs, and dual axes graphs.

When creating tables and graphs, it is always helpful to think of them as stand-alone documents. That is, if the table or graph you have created were to get copied or removed from the larger text or presentation it comes from, would it contain enough information for a user to make accurate inferences, know what the data represent, and know where the data are from. Consider, for example, Table 2-11.

This table presents information for a subset of patients at Port City Hospital who were overweight or obese upon admission by age group for 1 year. There are a number of ways to present this information graphically. These data are descriptive, in that we are measuring one thing, the percent of patients who were overweight or obese upon admission, and then segmenting the data by the age of the patient. Ages have been broken into categories for easier presentation, which is often done with continuous data. If we had attempted to count how many patients fell into each age, theoretically ages 1 through 100 plus, we would have over 100 bars on our chart. Therefore, categories help to distill down the data. For data arranged into categories, bar or column charts should be used. They can then depict either the

Table 2-11 Percent of Patients Overweight or Obese by BMI Score

Port City Hospital, 2008

Age group	Percent	(95% CI)	Sample Size (*n*)
18–24	36.7	(30.2, 43.3)	93
25–34	48.6	(44.3, 53.0)	289
35–44	56.6	(53.2, 60.1)	519
45–54	65.3	(61.7, 69.0)	488
55–64	68.1	(63.8, 72.3)	353
65 and older	59.7	(55.6, 63.8)	389

BMI $\geq$ 25 is considered overweight, as BMI $\geq$ 30 is considered obese.

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

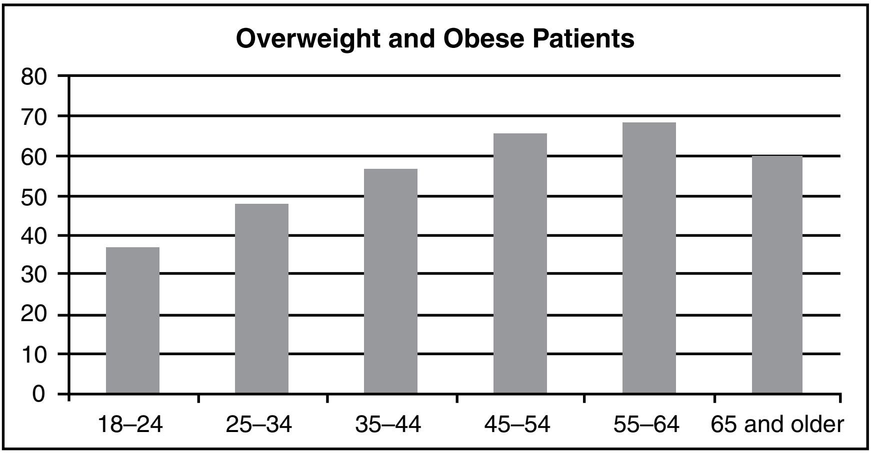

Figure 2-4 Overweight and Obese Patients

number of percent of the data points that fall into each category. Figure 2-4 shows one graphical representation.

Note here that we have bars for each age category and a title that states "Overweight and Obese Patients." Although semi-informative, there are limits to what we know about these data. If this chart were given to you by itself, what other questions might you have about its contents?

Figure 2-5 shows a slightly different view. Here error bars have been added to represent the 95% confidence intervals from column three of Table 2-11.

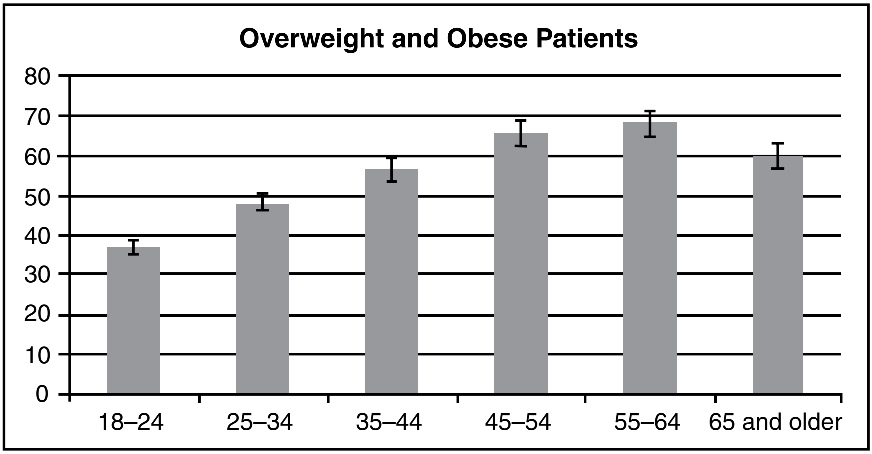

Figure 2-5 Overweight and Obese Patients Part 2

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

These are important because they denote that the data are taken from a sample of patients and should not be considered as representing the entire patient population. They also show that there is a 95% probability that the "true" population percentile for overweight and obese patients fall between the two ends of each error bar.

Finally, in Figure 2-6 we include the source of the data (Port City Hospital for 2007), labels for the *y*-axis, and a definitional text box for what BMI levels constitute overweight and obese. Figure 2-6 does the best job at becoming a stand-alone graphic, and conveys a detailed level of information.

A second form of categorical presentation is the pie chart. A perennial favorite with students, the pie chart should only be used when the categories constitute parts of a whole. Consider for example Table 2-12 and Figure 2-7. The slices of the graph represent divisions of the entire medical/surgical expenditures for the hospital and are exhaustive. When using pie charting there are a number of formatting considerations. The first is whether or not to include data labels, and then whether to use percents or whole numbers when labels are used. Figure 2-5 provides an example of using percents; however, it may be more appropriate to include the dollar amount each category represents, or perhaps both. The use of the data and their audience should define which is used. A final consideration is whether or not the graph will be printed in color. Colors allow for a clearer distinction between slices, whereas black and white graphs necessitate the use

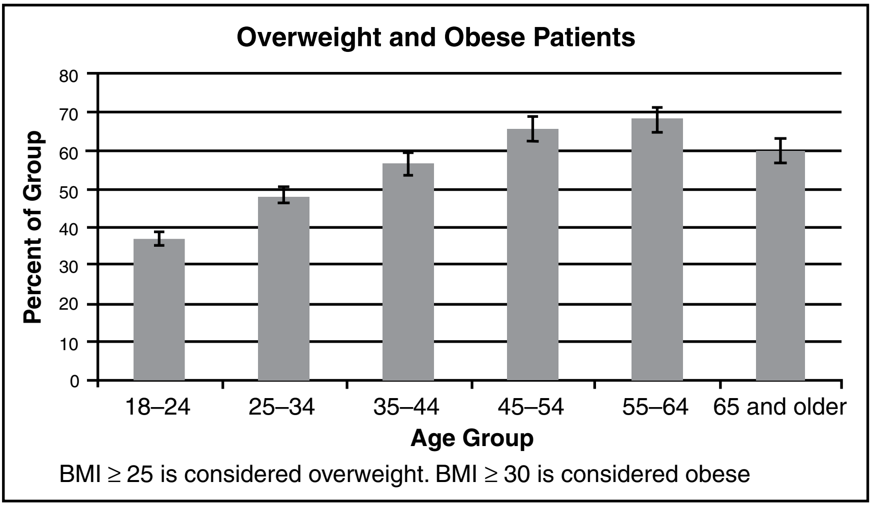

Figure 2-6 Overweight and Obese Patients: Port City Hospital, 2007

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

Table 2-12 2007 Expenditure Categories

Port City Hospital	*Med Surg*	*ICU*
Supplies	$ 189,654.00	$ 210,157.00
Professional	$ 1,085,623.00	$ 1,527,560.00
Pharmacy	$ 228,290.00	$ 142,152.00
Ancillary	$ 45,620.00	$ 33,158.00
Facilities	$ 624,877.00	$ 218,906.00
Administrative	$ 328,176.00	$ 3,235,148.00
Total	$ 2,502,240.00	$ 5,367,081.00

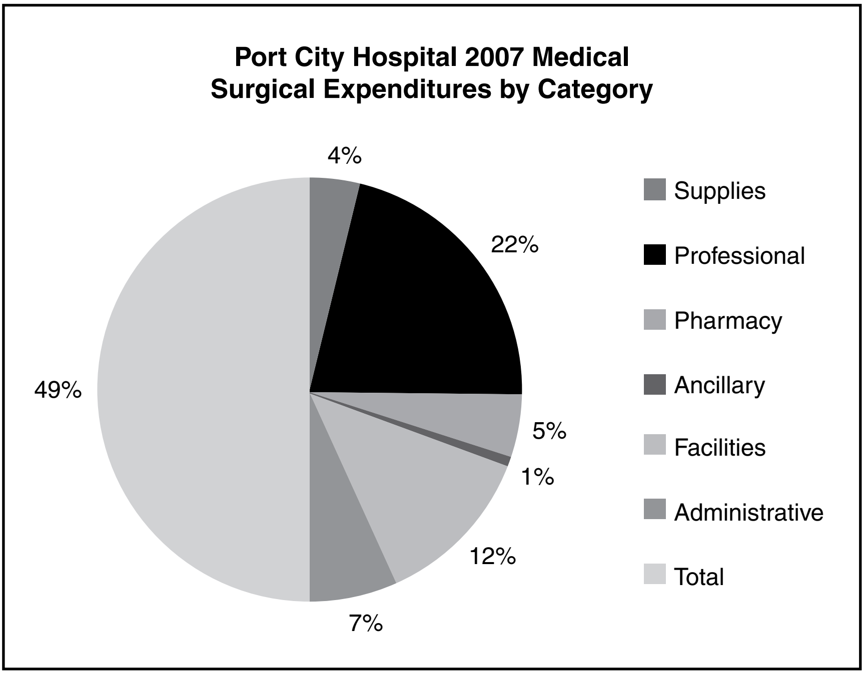

Figure 2-7 Port City Hospital 2007 Medical Surgical Expenditures by Category

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

of patterns within the slices to differentiate between them. Care should be taken to ensure that no two categories are easily confused because of similar hues or patterns.

A third basic technique for visually displaying information relates to time series data. When observations are measured over time, line graphs should be used to reflect the continuous nature of the data. Line graphs also allow for examination of visual trends over time that may not be readily discernible in table form. Consider Table 2-13, which examines births at Port City Hospital from 2005 through 2008. The data as presented do not lend to the immediate identification of trends by month, collection of months, or year. However, when placed in graphical form, as shown in Figure 2-8, two trends become evident. The first is that the number of births has risen each year. The second is that there a constant trend is reflected throughout the year shown by the peaks and valleys in the data over time. This seems to reflect a seasonal trend in which births spike in May and then decline during summer months, rising again late in the year. By examining these data using a stacked line graph, we can see both trends simultaneously.

The final graphical tool we will explore is the dual axis graph. This type of graph is suitable when two sets of data are to be shown together on one graph,

Table 2-13 Port City Hospital Births, 2005–2008

	2005	*2006*	*2007*	*2008*
Jan	21	25	30	39
Feb	25	30	35	42
Mar	21	31	37	51
Apr	24	34	41	53
May	35	35	42	57
Jun	14	20	30	44
Jul	21	23	27	41
Aug	27	25	31	40
Sep	33	37	45	55
Oct	37	40	50	60
Nov	30	38	42	62
Dec	36	45	48	58

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

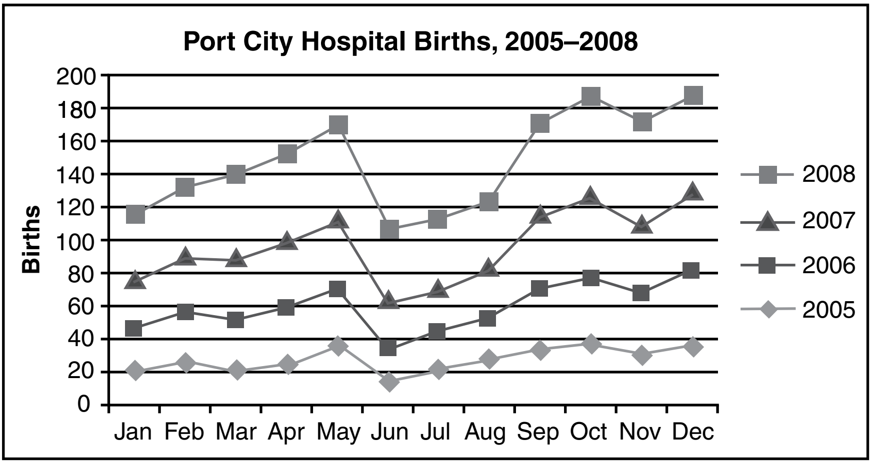

Figure 2-8 Port City Hospital Births, 2005–2008

but the units of measurement of the two variables are different. For example, consider Table 2-14, which shows the number of full time equivalent (FTE) employees and the total expenditures by hospital department. The visual representation allows for the comparison of two data elements, one measured in person equivalents and one in dollars to be viewed simultaneously (Figure 2-9).

Table 2-14 FTE Employees and Total Expenditures by Department

Port City Hospital 2008		
	FTE employees	*Total Expenditures*
Med/Surg	23	$ 5,645,230.00
ED	14	$ 825,180.00
ICU	17	$ 1,236,450.00
Neonatal	12	$ 1,647,264.00
Radiology	6	$ 546,230.00
Lab	6	$ 427,451.00

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

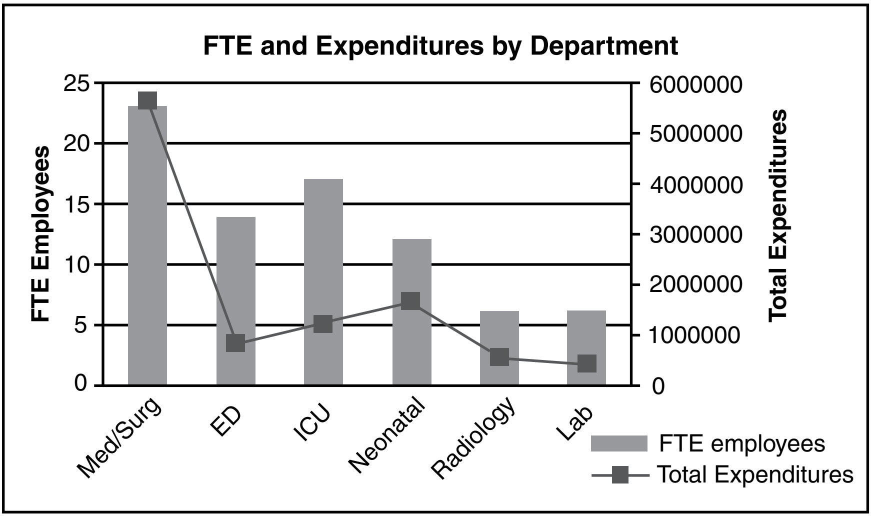

Figure 2-9 FTE and Expenditures by Department

It would be incorrect to place them relative to one vertical axis given that FTEs are measured in the tens (10, 12, 20 etc.), and total expenditures reach into the millions. To accommodate such large numbers for expenditures, the scale of the axis would reach up to the highest data point, here over $5 million, which would in turn mute any differences between the number of FTEs when those differences are so small.

Sound statistics and presentation skills afford the healthcare manager a clear and defensible way to both ask pertinent questions and construct evidence to answer those questions. The job of the competent manager is to know what the right questions are to ask, and then how to construct hypotheses and analyze data in an appropriate way. By following the simple rules presented here, data can be categorized by how they are measured (categorical or continuous) and following from that are a defined set of analysis options for each type. It is further incumbent upon the analyst to present the data in a clear and understandable way, but also to take care to describe fully and cite the data used. These basic analyses will provide the foundation for many of the other managerial competencies described later in this book.

EXERCISES

2-1 Your organization collects data on individual patients shown in Appendix Table 2. Identify whether each variable is measured nominally, ordinally, or as an interval/ratio variable.

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

2-2 What statistical measures would you use to summarize the variable for age? What about for gender? For convenience satisfaction? How would you present these graphically?
2-3 If you are interested in whether satisfaction scoring differed by the amount the individual paid as a co-pay, how would you state this inquiry as testable hypotheses (the null and alternative)? What statistical test would you run to test this hypothesis?
2-4 Perform the test defined in Table 2-3. State your conclusions.
2-5 Compare gender and having Rx coverage. State your hypotheses. Perform the appropriate statistical test and interpret.

Appendix Table 1 Chi-Square Values by Alpha Level and Degrees of Freedom

Df	*0.5*	*0.1*	*0.05*	*0.02*	*0.01*	*0.001*
1	0.455	2.706	3.841	5.412	6.635	10.827
2	1.386	4.605	5.991	7.824	9.210	13.815
3	2.366	6.251	7.815	9.837	11.345	16.268
4	3.357	7.779	9.488	11.668	13.277	18.465
5	4.351	9.236	11.070	13.388	15.086	20.510

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

Appendix Table 2 Patient Satisfaction, by Insurer, Same Day of Appointment, Prescription Coverage and Level of Co-pay

Gender	*Age*	*Convenience Satisfaction*	*Insurer*	*Same day appointment?*	*Rx coverage*	*Co-pay ($)*
M	22	5	Select	Y	Y	15
M	24	4	Select	Y	N	15
F	45	5	Select	Y	Y	10
F	38	5	Select	Y	Y	10
F	48	3	Select	Y	N	10
F	50	4	Select	Y	Y	20
M	67	4	Medicare	Y	Y	5
F	23	5	Tri-state	N	Y	10
F	19	5	Tri-state	N	Y	15
M	14	3	Tri-state	Y	Y	20
F	27	5	Reliant	Y	N	15
M	33	3	Reliant	N	N	0
F	39	4	Reliant	N	N	5
F	47	4	Tri-state	N	Y	5
M	42	5	Tri-state	N	N	5
M	31	5	Reliant	N	N	15
M	20	4	Reliant	N	Y	15
F	72	5	Medicare	Y	Y	15

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION

F	44	4	Tri-state	N	N	20
M	45	5	Select	Y	N	20
F	60	3	Reliant	Y	Y	20
M	63	5	Tri-state	Y	Y	5
F	27	5	Tri-state	Y	Y	5
M	68	5	Medicare	N	Y	5
F	33	5	Reliant	N	Y	10
F	38	3	Reliant	N	Y	15
F	55	4	Medicare	N	Y	15
M	51	5	Select	Y	Y	15
F	48	4	Tri-state	Y	Y	20
M	49	5	Tri-state	Y	Y	0
M	55	5	Select	N	Y	5
F	61	4	Reliant	Y	N	0
F	23	5	Medicare	Y	N	0
M	69	5	Medicare	N	Y	20
F	41	4	Tri-state	Y	N	5
F	14	4	Medicare	Y	N	0

© Jones and Bartlett Publishers, LLC. NOT FOR SALE OR DISTRIBUTION