Section III

Monitoring Performance

Five
CHAPTER **5**

Performance Monitoring Basics

Here is an example to illustrate the need of monitoring. You want a mortgage loan with a 6.0% interest, but the bank insists on a credit score of 750. Your score is 690, which would get a loan with a 7.0% interest rate. If you want a better rate, you need a better credit score. That 1% difference in the interest rate would cost you more than $70,000 dollars in payment for a $300,000 loan over 30 years.

What Is Performance Monitoring and Why Do It?

Why keep tracking your credit score? First, it provides an ongoing check on your financial performance records. You may find unusual financial irregularities and misdeeds such as signs of identity thefts or mistakes in your financial transaction records. Second, it helps you uncover inefficient practices and operations that lead to poor finance performance. You may discover that your low credit score is a result of overspending and late payments of credit card bills. To improve the score, you need to cut back the spending level and pay the debts on time. Third, the monitoring also helps you establish the need for services. For example, after reviewing the credit report, you may realize that you need credit consulting and financial planning to develop a healthy financial life and to achieve your financial goals.

Finally, perhaps most important, the monitoring helps you uncover underperformance quickly to avoid the further deterioration of the performance. Imagine that your low credit score is caused by the tardiness to pay credit card bills, and this tardiness goes unnoticed long enough for you to have trouble getting any credit lines. The result may be declamation of insolvency for your personal finance that directly and significantly affects your life.

Performance monitoring is a systematic and periodic observation of performance over time in order to develop or verify performance records, to uncover inefficient and ineffective practices, to identify needs for services, and most important, to detect underperformance timely to avoid the further deterioration of performance. It is a systematic tracking and ongoing examination of an organization's performance by weighing it against established performance standards.

Performance monitoring is different from once-in-a-while performance auditing or evaluation in that it is conducted more frequently, and it is more focused on tracking daily operations. Performance monitoring is aimed at identifying the symptom of underperformance quickly and responding timely.

Performance monitoring should apply to all services in public and nonprofit organizations. Some services need it more than others because the underperformance of these services needs swift and timely responses, and a delay in response could have devastating consequences. Crimes need to be observed frequently by the police to detect criminal patterns and to develop proper strategies to save lives and properties. Emergency supply during natural or man-made disasters should be tracked closely to avoid the shortages that could cause the loss of lives. Traffic accidents need to be watched carefully to minimize fatalities and loss of properties. Response times to emergencies should be monitored closely to avoid or to reduce the loss of lives. The availability of cash to pay off financial obligations should be tracked frequently to prevent financial insolvency. This chapter introduces the steps in performance monitoring and basic monitoring tools. Chapters 6 and 7 present two specific monitoring tools.

Developing a Performance Monitoring System

The steps in a performance analysis in Chapter 2 apply to the development of a performance monitoring system: understanding the issue for monitoring, asking the right monitoring questions, developing a theory for monitoring, developing measures for monitoring, determining data collection methods, collecting data for monitoring, conducting performance monitoring through analyzing data with monitoring tools, and writing the monitoring report. These sequential steps are shown in Figure 5–1.

Understanding the Issue for Monitoring

Imagine that a police department has recently received an increasing number of citizen calls for police assistance and wants to know if this is a seasonal surge due to the temporary influx of visitors for the college spring break or a more permanent increase in criminal activities. Imagine that an emergency response unit piles up food and water for hurricanes only to find out what people really need is medicines and medical supplies. Imagine that the financial director in your city has just told you that the city lacks cash to pay for its bills because a few large taxpayers have not paid their shares of taxes on time.

The very first step in developing a performance monitoring system is to identify a *monitoring need*. It is quite common that the need starts with an issue in everyday management operations and decision making. Increasing citizen complaints about tardiness in garbage collection indicate a need to monitor the garbage collection operation. More frequent requests for

Figure 5–1

Performance Monitoring Steps.

Understanding the issue for monitoring → Asking questions → Developing monitoring theories → Developing measures for monitoring → Determining data collection methods → Collecting data → Conducting performance monitoring → Writing the monitoring report police assistance call for a possible monitoring of crime activities. A rising number of traffic accidents suggest a need to monitor and improve traffic control. The appearance of large and constant budget deficits requires the monitoring of revenue collection efforts and spending patterns.

A clear monitoring need helps you determine the *monitoring goal(s)*: what you want to achieve in the monitoring. Stated above, performance monitoring can help you develop and verify performance records, discover inefficient or ineffective practices and operations, identify service needs, and/or uncover underperformance quickly to avoid further deterioration of performance. A performance monitoring often has multiple goals.

For your specific monitoring issue, the goal can be further specified as, for example, to meet residents' needs for emergency supplies during disasters (identification of service needs); to increase the chances of timely garbage collection (timely discovery of inefficient or ineffective operations and underperformance); to improve police response to crimes (timely discovery of underperformance); to reduce traffic accidents in rush hours (timely identification of service need patterns, discovering inefficient practices, and uncovering underperformance); or to reduce budget deficits or to balance the budget (discovery of inefficient practices).

A good understanding of the monitoring need and goals helps you accurately determine what should be monitored—*the monitoring subject(s)*. Should you track organizational inputs, outputs, or outcomes? Should you monitor revenues and manpower consumed for the production (efficiency) or for goals achieved (effectiveness)? Determination of a monitoring subject will be discussed again later in this chapter when the key components of a monitoring flow and monitoring measures are discussed.

Determining Monitoring Questions

A good understanding of the monitoring need and goal should help the development of the *monitoring questions*. Facing a growth in citizens' calls for services and a goal to improve responses, a police department may want to ask the following questions: Is the growth unusual, compared with data in the past? Has there been an increase in crimes committed since last year? Has there been an increase in crime-fighting activities in the police department? With an increase in traffic accidents and a goal to reduce traffic congestion, a transportation authority may want to know the following: Is the recent increase in traffic fatalities unusual? Has there been an increase in traffic volume? Has the city's plan to ease traffic congestion been implemented properly and timely?

Although every monitoring has its own specific questions, the following is a list of generic forms of performance monitoring questions. They should help you develop your own monitoring questions.

- Are performance goals being met?
- Has the performance plan been implemented effectively?
- Have operations been implemented according to the plan?
- Are the intended services being delivered to the intended clients?
- How good is my performance compared with others' performances, my previous performance, and the performance standard?

- Has the agency performed better or worse? Has the agency underperformed?
- Are there any signs of underperformance?
- Is there any room for performance improvement?
- Is my performance unusually poor, compared with data in the past?

Developing a Theory of the Monitoring Flow

Say you are a police chief in a city government who has heard many citizen complaints about crimes recently, and you want to know if there has been a significant increase in criminal activities in your jurisdiction. To respond swiftly to any surge of crimes, you want a system to monitor police response to crimes. You articulate that the monitoring goal of the system is to improve police response to crimes. The monitoring subject is police response to crimes. What is it? How do you measure it?

One way to understand police response is to specify its role in the production process discussed in Figure 1–1 of Chapter 1 (Inputs → Process → Outputs → Outcomes). The inputs are resources consumed in providing police response, which can be the budget amount allocated for the police patrol or the manpower designated for police response. The process concerns the activities and operations in providing police response. The outputs are police responses directly generated from the process such as arrests, while the outcomes are the results expected from police response such as crime reduction. This flow of inputs to outcomes in performance monitoring is the *monitoring flow*, shown in Figure 5–2. Monitoring subjects can be inputs, outputs, process, and/or outcomes in the monitoring flow.

Thus, a well-developed theory of the monitoring flow should help you specify the monitoring subject and its role in the production process—a critical step for developing proper monitoring measures and for applying the monitoring results to address any underperformance in the production process. It is important to note that a monitoring system should emphasize the monitoring of outputs and outcomes, although the system may also track inputs and the process that lead to the output and the outcome. Monitoring inputs and the process may provide clues on how to develop proper strategies to improve the output and outcomes.

Developing Measures for Monitoring

Once the monitoring flow and monitoring subjects are determined, monitoring measures should be developed accordingly. Table 5–1 lists possible measures in the police response example, along with their monitoring subjects and their roles in the production phases.

Figure 5–2

The Monitoring Flow: A Case of the Production Process of Police Response.

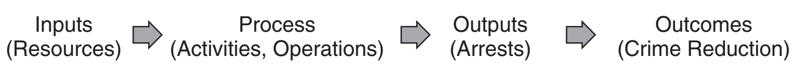


Table 5–1

Developing Measures for Monitoring Police Response

Production phase	Inputs in police response
Monitoring subject	Resources consumed in police response
Measure example	The number of sworn officers on duty per 1000 populations
Production phase	Process in producing police response
Monitoring subject	Activities and operations in police response
Measure example	The number of calls for services dispatched
Production phase	Outputs of police response
Monitoring subject	Products of police response generated from the process
Measure example	The number of arrests made
Production phase	Outcomes of police response
Monitoring subject	Desirable results of police response
Measure example	The reported number of violent crimes per 1000 populations

What measures should be used in monitoring? Because numerous measures are available for each possible monitoring subject, using all of them would be too costly and time consuming. Selecting limited measures is necessary. In addition to the criteria discussed in the "Developing Performance Measures" section in Chapter 2 (e.g., measurement validity, measurement reliability), consider the following in the selection.

First, measures selected should meet monitoring goals. To ensure swift police response to crimes, the response time and crime rates may be monitored. To discover residents' needs for a police service, the number of calls for the service may be monitored. To ensure availability of sufficient resources to support police response, the police budget amount may be monitored. If you are concerned about the possibility of resource waste in operations, measures of efficiency of police operations, such as the number of arrests per police employee, may be monitored.

Second, indicators selected should address specific monitoring needs. For example, an increase in users' complaints for water and sewer services may indicate a need to improve user services. Monitoring the number of user complaints may become necessary. A surge in traffic accidents may suggest a need to improve the transportation system, so monitoring the measures of transportation practices such as the miles of newly paved roads and the miles of existing roads widened or maintained in good condition may be needed.

Collecting Data for Monitoring

Data for performance monitoring come from surveys, archival documents or data, or interviews of related persons. A key component in data collection is to determine *monitoring frequency*—how often a measure should be monitored. Monitoring frequency dictates how often performance data should be collected. In the above case of police response, how frequently should we monitor the number of calls, the number of arrests, or the crime rate statistics: every day, every month, every 2 months, or even longer? Monitoring goals determine monitoring frequency. Daily monitoring of data is unnecessary for a monitoring intended to produce police response information for an annual budget request; monthly collection of the data may be enough. Monthly or annual data should be enough to demonstrate the trend of police response for planning or legislative oversight. Similarly, daily or monthly monitoring may be unnecessary for constructing a 5-year strategic plan. Nonetheless, if the monitoring is for improving day-to-day police operations, daily monitoring and collection of the response information are needed.

The monitoring cost is another consideration in determining monitoring frequency. It could be costly in money and manpower to conduct daily or even monthly monitoring. The cost incurred in frequent monitoring could exceed the benefits of the information yielded in the process.

Conducting Performance Monitoring and Writing the Monitoring Report

Conducting performance monitoring is to use *performance monitoring tools* to discover service need patterns, inefficient–ineffective practices, and underperformances. The essence of these tools is a systematic comparison of a performance with established performance standards, performance variations, or standardized performance. Thus, performance monitoring tools can be classified into three categories: tools in monitoring against performance standards, tools in monitoring performance variation, and tools in monitoring standardized performance. This chapter presents basic monitoring tools, while Chapters 6 and 7 cover two advanced tools of monitoring.

Once the monitoring is completed, the results should be presented in a monitoring report that follows the format discussed in Chapter 2 (“Writing a Performance Analysis Report”). It is a good practice to write a monitoring report for the record even if no warning trend of underperformance is detected.

Monitoring Against Performance Standards

Let us continue our police response example. Say that you, the police chief, want a system to monitor police response to crimes. You want to respond swiftly to any surge of crime activities in your city, articulating that the monitoring goal of the system is to improve police response to crimes, the monitoring subject is the outcome of police response, and the monitoring question is has there been an unusual increase in crimes recently.

You specify the outcome of police response by developing a monitoring flow like that described in Figure 5–2, selecting measures like those in Table 5–1. One outcome measure is the crime rate, the reported violent crimes against people (number of crimes per 1000 populations). Table 5–2 presents the crime rate for the past 2 years up to October of this year. Monthly data are monitored for timely police response.

The core analysis in performance monitoring is a process of using monitoring tools to consistently and systematically conduct performance comparison. One basic tool is the *comparison with the past performance* in which the current performance is weighed against

Table 5–2

Number of Reported Violent Crimes Against People per 1000 Populations

Monitoring Goal: To improve police response to crime
Monitoring Subject: Crime-responding outcomes
Monitoring Question: Has there been an unusual increase in crimes recently?
Outcome Measure 1: Reported violent crimes against people (number of crimes per 1000 populations)
Reporting Date: November 1, this year

	Monthly this year (column 1)	Monthly last year (column 2)
January	0.2120	0.1753
February	0.1541	0.2568
March	0.2561	0.2624
April	0.3184	0.3115
May	0.3154	0.3762
June	0.3306	0.2950
July	0.3863	0.3444
August	0.2172	0.1438
September	0.3110	0.2558
October	0.1619	0.1954
November	Not Available (or NA)	0.1429
December	NA	0.2326
Annual		2.9920

the past performance in order to discover any performance difference like that shown in Table 5–3.

Monthly crimes last year and this year until October are put side by side and analyzed. The data indicate that the crimes were relatively low from January to March, gradually increased and reached the highest during the summer months from April to July, then started to decline from August.

Also shown in Table 5–3 is the month-to-month comparison of the crimes. Of the 10 first months this year, 6 had a crime increase and 4 saw a decline from the same months last year. The data show a clear trend of monthly increase during the summer months from June to September, when the largest monthly increase was in August from 0.1438 last year to 0.2172 this year, for a stunning 0.0734/0.1438 = 51.0% increase! Also, the September increase of 0.0552, the second highest monthly surge for a 0.0552/0.2558 = 21.6% increase, is also really alarming. The monthly comparison shows a clear trend of crime increase in the summer, which constitutes a *warning sign of underperformance* of police response.

Table 5–3

Crime Rate: Comparison with the Past Performance

	Monthly this year (column 1)	Monthly last year (column 2)	Monthly difference (column 3 = column 1 − column 2)
January	0.2120	0.1753	0.0366
February	0.1541	0.2568	–0.1027
March	0.2561	0.2624	–0.0062
April	0.3184	0.3115	0.0069
May	0.3154	0.3762	–0.0607
June	0.3306	0.2950	0.0357
July	0.3863	0.3444	0.0419
August	0.2172	0.1438	0.0734
September	0.3110	0.2558	0.0552
October	0.1619	0.1954	–0.0336
November	NA	0.1429	NA
December	NA	0.2326	NA
Annual		2.9920	

In the above example, we compare crimes of the same months for the past 2 years. This same-period comparison is important in crime rate monitoring because crimes often show seasonal differences. Other examples of performances associated with seasonal fluctuations include unemployment rates, revenues generated from economic activities (such as sales taxes), traffic volumes, and attendances of park and recreation facilities. The same-month comparison is important in monthly monitoring.

In addition to the same-period comparison, you can also compare the performances of 2 consecutive months. The crime rate this February was 0.1541, a significant decline from that in the previous month (0.2120). Because the crimes tend to gradually increase from the start of a year, this large February decline indicates an anomaly in the crime pattern. In performance analysis, a performance data point that unusually departs from the existing data pattern is known as a *performance outlier* (see also Chapter 3). Outliers often occur due to particular events, which could be that the city's weather was particularly cold this February and that led to fewer criminal activities on the street, the city intensified its efforts in patrols and arrests in that month that drove criminals out of the city, or a change in the crime reporting method resulted in some crimes previously reported in February are now reported in January or March.

The examination of outliers helps discover these particular events that cause performance change and thus provides insights about the performance. However, one also should not read too much into outliers in performance monitoring because outliers, by definition, are unlikely to happen again. For example, if it is known that the remarkable low crime rate in February was a result of some specific events that are unlikely to happen frequently in the future, there should be little concern about the large crime rate difference between February and March (0.2561 − 0.1541 = 0.102 or about 10 more crimes per 10,000 populations).

One statistical way to alleviate the impact of outliers in performance monitoring is the *comparison with the average performance*. In the police response example, you can compare the March crime rate this year (0.2561) with the average of the first 3 months of the year [(0.2120 + 0.1541 + 0.2561)/3 = 0.2074]. Table 5–4 shows the comparison between a monthly crime rate and the average crime rate of the months so far this year. For example, the average in February is 0.1830 [(0.2120 in January + 0.1541 in February)/2]) in the second column of Table 5–4. Notice that the average performance can be replaced or complemented by the median performance or the mode performance.

Shown in column 3 in Table 5–4 of the 10 months in the comparison, the crimes are lower than the average in 3 months and equal to or above the average in 7 months. The comparison shows a trend of crimes higher than the average from March to September except a slight decline in August, which is indicated by all positive figures in column 3 from March to September except that in August, constituting a sign of underperformance during this period.

In addition to comparing with the past and the average, you can measure the performance up against the national, the state, or the regional standards or benchmarks. This is called the *comparison with established performance standards* (or *benchmarks*). The benchmarks could be national, state, or regional averages; the performance of similar organizations; or any established or acceptable performance standards. In the police response example in Table 5–5, the monthly performance is compared with the last year's average crime rate (0.2767) in the cities of similar population sizes in the same geographic region. The 0.2767 is the annualized monthly average obtained by dividing the annual rate by 12 (the actual monthly data of the regional average is not available). The monthly crime rate is also compared with last year's national monthly average (0.3445).

Table 5–4

Crime Rate: Comparison with the Average Performance

	Monthly this year (column 1)	Mean this year (column 2)	Difference from the mean (column 3 = column 1 – column 2)
January	0.2120	0.2120	0.0000
February	0.1541	0.1830	–0.0290
March	0.2561	0.2074	0.0487
April	0.3184	0.2351	0.0833
May	0.3154	0.2512	0.0642
June	0.3306	0.2644	0.0662
July	0.3863	0.2818	0.1045
August	0.2172	0.2738	–0.0566
September	0.3110	0.2779	0.0331
October	0.1619	0.2663	–0.1044
November	NA		
December	NA		
Annual			

Table 5–5

Crime Rate: Comparison with Performance Standards

	Monthly this year (column 1)	Monthly regional average (last year) (column 2)	Difference from regional average (column 3 = column 1 – column 2)	Monthly national average (last year) (column 4)	Difference from national average (column 5 = column 1 – column 4)
January	0.2120	0.2767	–0.0647	0.3445	–0.1325
February	0.1541	0.2767	–0.1226	0.3445	–0.1904
March	0.2561	0.2767	–0.0206	0.3445	–0.0884
April	0.3184	0.2767	0.0417	0.3445	–0.0261
May	0.3154	0.2767	0.0387	0.3445	–0.0291
June	0.3306	0.2767	0.0540	0.3445	–0.0139
July	0.3863	0.2767	0.1097	0.3445	0.0418
August	0.2172	0.2767	–0.0595	0.3445	–0.1273
September	0.3110	0.2767	0.0344	0.3445	–0.0335
October	0.1619	0.2767	–0.1148	0.3445	–0.1826
November	NA	0.2767	NA	0.3445	NA
December	NA	0.2767	NA	0.3445	NA
Annual		3.3200		4.1340	

The results indicate that the city's crime rate is higher than the regional average in 5 of the 10 months. Nonetheless, except in the month of July, the city's crime rate is lower than the national average. The city's crimes were higher than the regional average from April through September except August, suggesting a sign of underperformance during this period. This result is consistent with that from the above comparison with the average performance in Table 5–4.

In sum, the current performance can be compared against the past performance, the average performance, and the established performance standards in performance monitoring. These comparisons apply to all forms of performance measures including percentage measures, aggregate measures, average measures, and per capita measures as well as trend measures of growth rate, percentage difference, and percentage ratio.

Monitoring Performance Variation

Table 5–6 shows performances of 2 garbage collection teams by the tons of the garbage collected in the past 5 days. Which team performs better?

Measured by the average performances, both teams perform at the sample level—30 tons of collection a day. However, which team would you like to pick up your garbage? A closer look at the data shows that there is a larger day-to-day performance difference by Team A,

Table 5–6

Tons of Garbage Collected

Day	Team A	Team B
1	10.00	28.00
2	20.00	29.00
3	30.00	30.00
4	40.00	31.00
5	50.00	32.00
Mean	30.00	30.00

from 10 tons to 50 tons a day, while the difference by Team B is much smaller, from 28 to 32 tons a day, at around 30 tons a day. The performance of Team B is much more consistent and predictable.

Performance variation concerns the difference or the disparity of the performance. Statistically, it reflects the degree of the departure of individual performances (e.g., the ton of garbage collected each day) from the average performance. Monitoring performance variation reveals *performance predictability* and *performance consistency*. Statisticians also use the term *reliable* to describe a consistent and predictable performance. A performance is said to be reliable if the individual performances center closely around the average performance. A performance lacks reliability if individual performances scatter all over so they depart greatly from the average performance. As you can imagine, monitoring performance variation is important because you want to ensure reliable services delivered to residents or customers.

Here is another example. Table 5–7 shows the response times of 2 fire and rescue stations in an urban city. The response times (in minutes) of the recent 10 incidents for each station are presented. Which station performs better? Which station would you like to respond to your emergencies?

The average response time is the same for both stations, 7.4 minutes, but Station B appears to have a larger performance variation, while Station A's performance is more predictable and consistent around 7.4 minutes. Exactly how much variation does each station have? Statisticians use variance, standard deviation, and range to measure performance variation.

Variance (σ^2 or s^2)

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

Each of the individual performance cases ($i = 1,2,3 \ldots N$) is represented by X_i. In Table 5–7, for example, the cases for Station A are the 10 response times. The mean is μ, N is the number of cases, and Σ is the summation symbol where every element in it should be added. The above formula is used when you include *all* performance cases of the study subject. In statistics, a data set that includes all cases of a study subject is known as the *population* (the concepts of

Table 5–7

Response Times of Two Fire and Rescue Stations (in minutes)

Responses of last 10 incidents	Station A	Station B
1	8.00	3.00
2	7.00	12.00
3	8.00	7.00
4	6.00	4.00
5	6.00	4.00
6	9.00	6.00
7	7.00	3.00
8	9.00	9.00
9	8.00	11.00
10	6.00	15.00
Mean	7.40	7.40

population and sample will be discussed with more detail in Chapter 7). For example, if you study all 24,500 students in a university, these 24,500 students constitute a population for your study. However, if you merely draw a sample from the population, the sample variance is

$$s^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}.$$

Individual performance cases ($i = 1,2,3 \ldots n$) are represented by X_i in the sample, $\bar{X}$ is the sample mean, and n is the number of cases in the sample (while the capital letter N represents the number of cases in the population). Notice that the *variance* essentially estimates the distance of each individual performance case from the mean performance (i.e., $X - \bar{X}$). A smaller variance indicates that there is a smaller distance from each individual performance to the mean performance, or in other words, individual performances center more closely on the mean, resulting in a more predictable (close to the mean) and consistent (around the mean) performance.

In our example in Table 5–6, the mean of the daily collections for Team A is 30 [(10 + 20 + 30 + 40 + 50)/5]. The numerator of the formula for the sample variance, 1000 [$(10 - 30)^2 + (20 - 30)^2 + (30 - 30)^2 + (40 - 30)^2 + (50 - 30)^2$], is divided by the denominator 4 ($n - 1 = 5 - 1$) to arrive at a variance of 250. The variance for Team B's performance is 2.50, much smaller than Team A's variance, so Team B's performance is more predictable, consistent, reliable, and thus better than Team A's performance.

Standard Deviation (σ or s)

The *standard deviation* is simply the square root of the variance. The population standard deviation is symbolized by σ, while the sample standard deviation is denoted by s. In the above example, the sample standard deviations are 15.81 ($\sqrt{250}$) for Team A's performance and 1.58 ($\sqrt{2.5}$) for Team B's. The standard deviation has similar interpretation of the variance. A larger standard deviation indicates a larger performance variation. The standard deviation is used more often as a measure of variation than the variance.

Maximum, Minimum, and Range

You can identify the largest numerical value of the performance (maximum) and the smallest value (minimum). The difference between them (i.e., maximum – minimum) is called the *range*. In the above example of garbage collection, the maximum, the minimum, and the range are 50, 10, and 40 for Team A, and 32, 28, and 4 for Team B. Note that the range is different from the performance ranges defined in Chapter 3. The former is a statistical concept, and the latter is a management concept. There can be multiple performance ranges for a group of data, but there is only one range for the data.

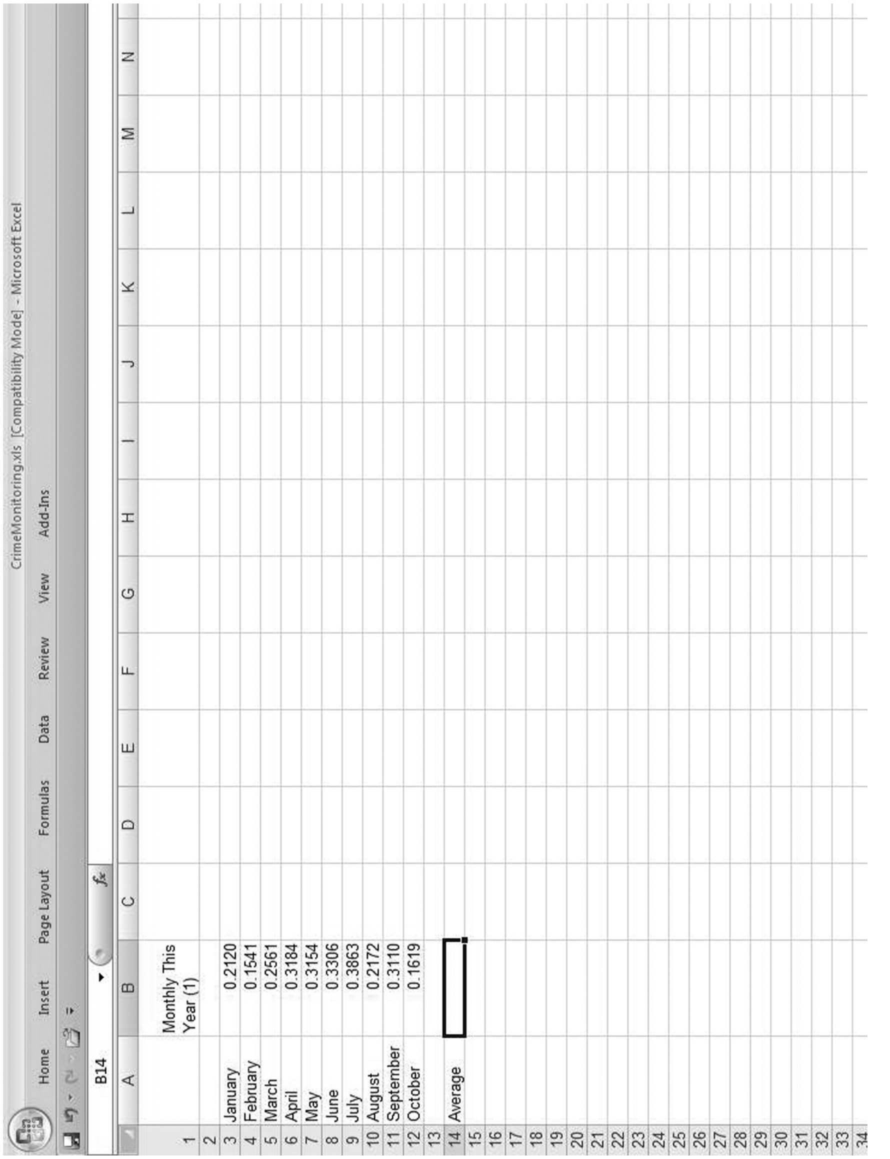
Excel Insert Function (f_x)

Chapter 3 introduces the Excel Data Analysis Descriptive Statistics procedure to obtain a group of descriptive measures such as the mean, the median, or the mode. If you want just one measure, the mean for example, Excel Insert Function (f_x) provides a much easier and more convenient way. For those who do not see the Insert Function symbol (f_x) in your Excel sheet, review Chapter 3 (section "Microsoft Office Excel") on how to load it. Let us say that you want to calculate the mean for the crime rate in Table 5–2.

1. Input the crime data in an Excel sheet, shown as the data array from B3 to B12 in Screen 5–1 and select a cell for the result of the average (B14 in this case).
2. Click the **Insert Function** button (f_x), next to the **Formula** bar, to open the **Insert Function** dialog box, shown in Screen 5–2.
3. Type the name of the function you want to use in the **Search for a Function** window in the **Insert Function** dialog box. If you do not know the name, select a category of formulas you want to work with from the **Select a Category** window. If you are not sure which category you should use, select the **All** category then click a function from the **Select a Function** window and read the definition of the function given below the window.
4. Select the **Average** function because you want to calculate the average of the data. Click **OK**.
5. Select data in cells from B3 to B12 (i.e., B3:B12) in the **Number 1** window (the only calculation in this example) in the **Function Arguments** box. You should be able to see the 0.2663 in the **Formula Result**, shown as in Screen 5–3.
6. Click **OK**. The result should appear in cell B14. You should also see the =AVERAGE (B3: B12) in the **Formula Bar**. In other words, you can use the Formula Bar directly to get your answer if you know the name of the function and the location of the data. Make sure you type the equal sign "=" each time you use the Formula Bar and place a bracket "( )" for the data.

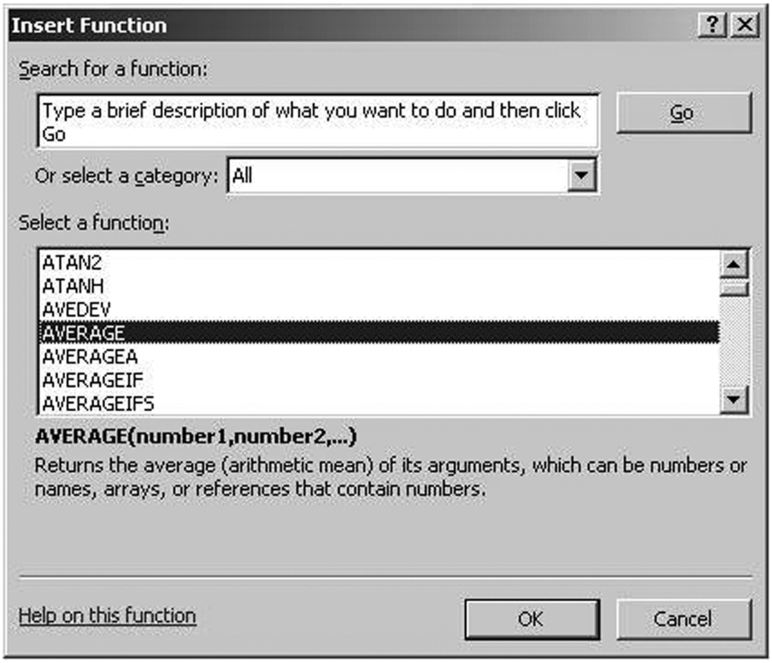
Screen 5–1

Excel Insert Function.



Screen 5–2

Insert Function Dialog Box.



The Excel Insert Function is a very useful calculation tool. Many calculations in this book use it. So knowing how to use it is very important to get through this book. For the convenience of the reader, Table 5–8 shows the names of some popular Insert Functions. A full list of Insert Functions used in this book is provided in Appendix A with related statistical topics covered.

Let us use the police response example to practice the above Excel Insert Function. Table 5–9 shows the measures of performance variation for crimes. For example, the variance so far this year in February (0.0017) is the variance of the crimes in the first 2 months (0.2120 in January and 0.1541 in February). The results show a steady increase over time in the variance and standard deviation, indicating that the crimes have fluctuated more over time, and they have become more unpredictable.

Monitoring Standardized Performance

Let us look at a simple example first. An academic department in a university uses both Graduate Record Exam (GRE) and Graduate Management Admission Test (GMAT) scores in the

Screen 5–3

Insert Function Argument Box.

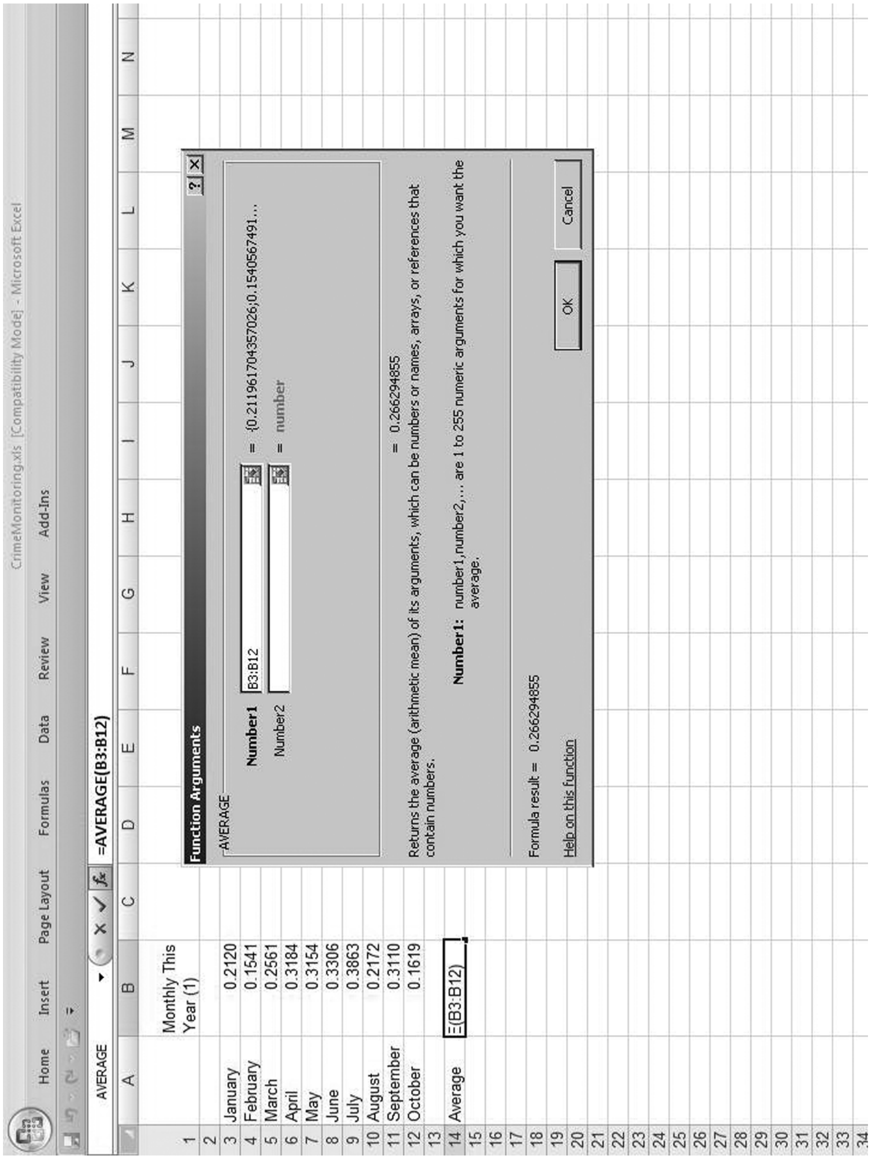


Table 5–8

Some Useful Excel Insert Functions (f_x)

Function	Name of function in Excel
Mean	AVERAGE
Median	MEDIAN
Mode	MODE
Population variance	VARP
Sample variance	VAR
Population standard deviation	STDEVP
Sample standard deviation	STDEV
Maximum	MAX
Minimum	MIN

Table 5–9

Crime Rate: Measures of Performance Variations

	Monthly this year (column 1)	Variance so far this year (column 2)	Standard deviation so far this year (column 3)	Maximum so far this year (column 4)	Minimum so far this year (column 5)	Range so far this year (column 6 = column 4 – column 5)
January	0.2120					
February	0.1541	0.0017	0.0409	0.2120	0.1541	0.0579
March	0.2561	0.0026	0.0512	0.2561	0.1541	0.1021
April	0.3184	0.0048	0.0695	0.3184	0.1541	0.1643
May	0.3154	0.0049	0.0701	0.3184	0.1541	0.1643
June	0.3306	0.0050	0.0706	0.3306	0.1541	0.1766
July	0.3863	0.0063	0.0792	0.3863	0.1541	0.2323
August	0.2172	0.0059	0.0768	0.3863	0.1541	0.2323
September	0.3110	0.0053	0.0729	0.3863	0.1541	0.2323
October	0.1619	0.0061	0.0779	0.3863	0.1541	0.2323
November	NA					
December	NA					
Annual						

graduate admission. Student A has a GRE of 1050, and Student B has a GMAT of 550, and the department can only accept one student. Which student should be accepted based on the test scores?

The GRE and GMAT scores cannot be compared directly because they use different scoring systems. GREs have a maximum score of 1600, while the maximum score in GMAT is only 800. So you have to use standardized scores for comparison. The standardized scores, also known as z-scores, put 2 different scores in the equal footing for comparison. The z-score of a population with a mean of μ and a standard deviation of σ is

$$z = \frac{X - \mu}{\sigma}.$$

Individual cases are represented by X. The z-score for a sample with a mean of $\bar{X}$ and a standard deviation of s is

$$z = \frac{X - \bar{X}}{s}.$$

In the above example, if the mean of GRE is 1000 with a standard deviation of 100, Student A's z-score is 0.5 [(1050 – 1000)/100]. If the mean of GMAT is 500 with a standard deviation of 50, the z-score of Student B is 1.0 [(550 – 500)/50]. Student B has a higher standardized score and therefore should be accepted by the department.

Standardization of performances allows the tracking and comparing of performances with different measures. For example, a police department can compare standardized crime rates with standardized response times to discover which measure shows a more severe sign of underperformance. Similarly, standardized scores make it possible to compare the performances of a fire department with that of a police department.

Table 5–10 shows the crime rate and the average response time in a police department for the past 10 months, from January to October. The z-scores of both measures are calculated

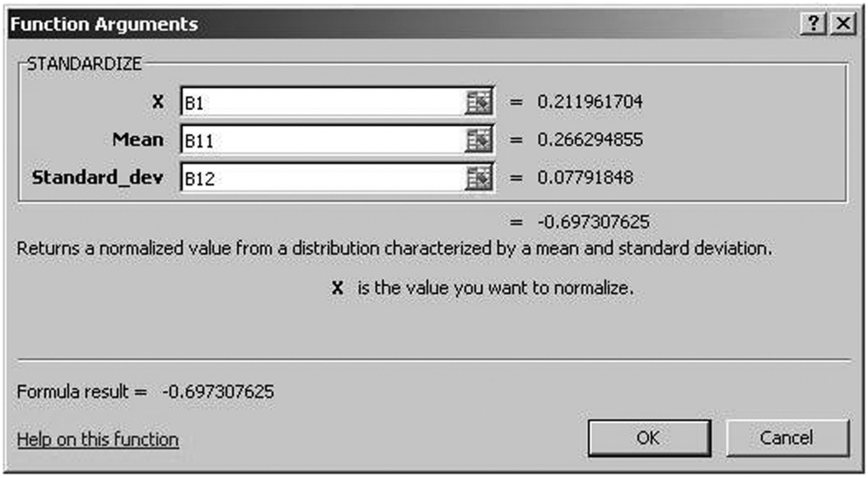
Table 5–10

Standardized Police Performances

	Number of crimes per 1000 populations	z-Score	Average response time	z-Score
January	0.2120	–0.6973	5.7000	–0.4675
February	0.1541	–1.4405	4.0000	–1.7536
March	0.2561	–0.1307	6.9000	0.4403
April	0.3184	0.6687	8.9000	1.9533
May	0.3154	0.6304	7.4000	0.8186
June	0.3306	0.8255	6.5800	0.1982
July	0.3863	1.5404	5.8000	–0.3919
August	0.2172	–0.6302	6.8000	0.3646
September	0.3110	0.5740	5.4000	–0.6945
October	0.1619	–1.3403	5.7000	–0.4675
Mean	0.2663		6.3180	
Standard Deviation	0.0779		1.3218	

Screen 5–4

Insert Function for *z*-Scores.



with the mean and standard deviation of the data. The result shows that both crime rates and the response time had best standardized performance in February when the crime rate was low (z = −1.4405), and the response time was short (z = −1.7536). Nonetheless, the police performance measured in the response time was better than that measured in the crime rate in February (–1.7536 is smaller than –1.4405). The worst police performance in the crime rate was in July (z = 1.5404), and the worst performance in the response time occurred in April (z = 1.9533), while the response time in April is worse than the crime rate in July from the comparison of their z-scores (i.e., 1.9533 is larger than 1.5404).

The *z-scores* can be easily calculated from the Excel Insert Function. The function name is Standardize. For example, the z-score of the crime rate in January (–0.6973) is calculated in Excel, shown in Screen 5–4. Notice that you can use Excel autofilling and absolute referencing introduced in Chapter 3 to generate z-scores of other months.

A Case Study

Many governmental services are outsourced to business or nonprofit contractors, creating a need to monitor the performance of the contractors to ensure the fulfillment of contractual requirements. The services likely to be contracted out at the local government level include, but are not limited to, air traffic control, legal services, fire protection, trash collection, health care, snow plowing, building maintenance, bill collection, data processing and analysis, street cleaning, steer repair, and recycling. The goals of contracting are usually high service quality and low service cost.

However, if poorly managed, outsourcing could go sour quickly. A government may fail to recognize all costs it should assume in outsourcing decisions. Also, outsourcing can cause a government's over reliance on external suppliers that results in a possible loss of control over critical activities and cost management. Most importantly, a government may lose control over the quality of service provided by a contractor.

The city of St. Stevenson (population 145,000) has used a private company to collect the waste of residents and businesses. Five years ago, the city signed a contract with the company to collect the solid waste and recycled materials from residents and businesses in the city. The contract specifies that the city should assign a full-time performance inspector to enforce the fulfillment of contractual terms with the company.

Step 1: Understanding the Issue

Mr. Eddie Jones is appointed as the performance inspector for the city. One of his very first responsibilities is to examine the contractual terms and to assess the need to regularly monitor the contractor's performance in the solid waste collection, the recycled materials collection, and the cleaning of the streets. He identifies several reasons that make the monitoring necessary.

First, waste collection is one of the most important services to the residents of the city. According to a recent citizen survey, street cleaning along with public safety are the important factors that influence residents' overall satisfaction toward the city government. Continual high performance in waste collection is expected by residents.

Second, waste collection contracts are often long term; contracts longer than 3 years are common. The contract in St. Stevenson was signed for 5 years. A long-term contract may result in less motivation and momentum for the company to improve services. Consistent monitoring may serve as a stimulus for the contractor to improve performance.

Third, waste collection in the city is a supplier market in which only two companies are qualified for and capable of providing the service for the city. Because of the limited competition, the cost to switch to another company if the current contractor performs poorly is very high. It saves taxpayers money if the city regularly monitors the performance of the current provider and provides swift feedback for performance improvement.

Bearing the monitoring need in mind, Eddie believes that the monitoring goals are (1) to frequently and systematically review the progress toward the completion of the contractual terms, (2) to identify areas of underperformance, and (3) to recommend proper performance improvement actions. He also determines that the monitoring subject should be the outcome of solid waste collection services, defined as the satisfaction of users (including residential and business users of the services) and a clean environment of streets, roads, and communities.

Eddie met with the city manager and the finance director to discuss the role of performance monitoring in management and to solicit support from them. He argued that contractual performance monitoring should be a critical part of the city's managerial control process, and the city should integrate the monitoring in its ongoing strategic planning and budgeting process, as well as in its performance measurement and reporting practice.

Step 2: Starting with Questions

Eddie's idea has strong support from the city manager and the finance director. The performance inspector position is secured and is specified as a full-time position to monitor the streets and collection services. Based on the monitoring goals, several monitoring questions have been developed by Eddie:

- Are the performance goals established for the waste collection met?
- How good is the performance compared with performance benchmarks?
- Are there any warning signs of underperformance?
- Is there any room for performance improvement?

Step 3: Developing a Theory for Monitoring

The success of the monitoring depends on the inspector's knowledge on the production process of the waste collection. Eddie must have an acute understanding of the production flow of waste management and the critical factors that influence the flow. He must use this knowledge to develop a proper measurement process that captures the resources used, the operational elements in which wastes are collected, the outputs of the process on how much waste is collected, and most important, the outcome in user satisfaction and clean environments.

Eddie should develop a theory on how the outcome is affected by the waste collection inputs, process, and outputs, as well as other nonperformance factors such as weather, which may affect the outcome of user satisfaction and clean environments. Using this theory, Eddie can identify the causes of an underperformance quickly and respond properly. He can discover whether the underperformance is *minor* (e.g., waste overflows from containers) that can be fixed quickly with a minor change of the process or *major* (e.g., constant delay in time for waste pickup) that requires structural adjustment of the production process.

Step 4: Developing Measures for Monitoring

The city has collected a total of 20 inputs, outputs, and outcome measures related to the waste collection services. These measures assess the inputs (e.g., the annual total expenses for waste collection, the number of full-time employees involved in waste collection contracting–monitoring–evaluation); the outputs (e.g., the average number of users served per month, the number of tons of waste collected, the percentage of collections made on schedule, the percentage of user complaints handled within 24 hours); and outcomes (e.g., the average response time to a user complaint for a follow-up, the daily number of user complaints, the percentage of complaining users who are satisfied with the follow-up—the user satisfaction rate).

Step 5: Collecting Data and Monitoring Performance

It is impossible to regularly monitor all 20 measures, so Eddie decides to monitor the 2 most important outcome measures: the daily number of user complaints and the percentage of complaining users who are satisfied with the follow-up. Both assess user satisfaction.

Eddie tracks down the number of complaints daily. The follow-up satisfaction rate data were assembled monthly, prior to his monthly meeting with the contractor. In practice, the majority of user complaint calls are received in the city and transferred to the inspector's office. Eddie contacts the contractor to convey the complaints daily by e-mail or phone. Once a complaint is resolved or an explanation is provided, the contractor informs Eddie, who then could call the users to conduct a follow-up survey of satisfaction. Table 5–11 shows the number of user complaints for the past 2 weeks.

Although Eddie communicates with the contractor daily to exchange information, he uses a 5-day period (from Monday to Friday) to decide on whether there is a sign of underperformance based on the number of complaints. He tallies and analyzes the following information every Friday in his monitoring practice:

- the daily number of complaints
- the daily average of this week vs. the daily average of this month until now
- the daily average of this week vs. the daily average of last month
- the daily average of this week vs. the daily average of last year

Comparisons are made for the total number of complaints and, more importantly, for the number of major complaints. From past experience, Eddie knows that a possible warning sign of underperformance is detected if one of the following occurs:

- An increase in the daily number of complaints 5 days in a row.
- An increase in the number of major complaints 5 days in a row .
- The daily average of complaints this week is larger than the daily average of complaints this month until now (or last month or last year).
- The daily average of major complaints this week is larger than the daily average of major complaints this month until now (or last month or last year).

If a warning sign is detected and Eddie finds the need for possible structural changes for the operation, he will call for a face-to-face meeting with the contractor to discuss potential changes needed for performance improvement. In this case, there were a daily average of 15.20 complaints for the week ending on April 6, higher than the daily average of March (14.60) but lower than the daily average of last year (15.32). Wednesday, April 4, saw the largest number of complaints. There were 8 major complaints during the week for a daily average of 1.60. Eddie has asked the contractor to provide a follow-up within a week on the corrective actions made on these major complaints.

There was a daily average of 10.80 complaints for the week ending on April 13, a decline compared with the daily average this month until now (13.00). Nonetheless, there were 9 major complaints this week instead of 8 last week. Eddie decides to raise his concern about this increase in the next meeting with the contractor.

Step 6: Preparing the Performance Monitoring Report

The key monitoring results are prepared by Eddie in a monthly report that is delivered to the city's management and the contractor. Eddie also prepares an annual monitoring report used in the city's annual budgeting process to make contracting and funding decisions.

Table 5–11

Number of User Complaints for Waste Collection in St. Stevenson in April

Date	Minor	Major	Total	Daily average this week (total)	Daily average this month until now (total)	Daily average last month (total)	Daily average last year (total)	Daily average this week (major)	Daily average this month until now (major)	Daily average last month (major)	Daily average last year (major)
Monday, April 2	11.00	3.00	14.00		14.00	14.60	15.32		3.00	3.30	3.60
Tuesday, April 3	17.00	3.00	20.00		17.00	14.60	15.32		3.00	3.30	3.60
Wednesday, April 4	25.00	2.00	27.00		20.33	14.60	15.32		2.67	3.30	3.60
Thursday, April 5	11.00	0.00	11.00		18.00	14.60	15.32		2.00	3.30	3.60
Friday, April 6	4.00	0.00	4.00	15.20	15.20	14.60	15.32	1.60	1.60	3.30	3.60
Monday, April 9	7.00	1.00	8.00		14.00	14.60	15.32		1.50	3.30	3.60
Tuesday, April 10	11.00	2.00	13.00		13.86	14.60	15.32		1.57	3.30	3.60
Wednesday, April 11	10.00	3.00	13.00		13.75	14.60	15.32		1.75	3.30	3.60
Thursday, April 12	10.00	1.00	11.00		13.44	14.60	15.32		1.67	3.30	3.60
Friday, April 13	7.00	2.00	9.00	10.80	13.00	14.60	15.32	1.80	1.70	3.30	3.60

Practices

Key Terms

Performance monitoring
Performance monitoring steps
Monitoring needs
Monitoring goals
Monitoring questions
Monitoring flow
Monitoring subjects
Monitoring frequency
Performance monitoring tools
Monitoring against performance standards
Comparison with the past performance
Comparison with the average performance
Comparison with established performance standards
Warning sign of underperformance
Monitoring performance variation
Performance variation
Performance predictability, consistency, and reliability
Variance
Standard deviation
Maximum, minimum, range
Excel Insert Function (f_x)
Monitoring standardized performance
z-Scores

Practice Problem 5–1

The department of transportation in a state government has four vehicle registration offices in a metropolitan area. The department recently conducted a series of examinations to find out the number of mistakes made in issuing or renewing driver licenses. Twelve tests were conducted in each of these four agencies, and the results are shown in Table 5–12. Input the data in an Excel file.

1. Calculate the mean, the sample variance, the sample standard deviation, the maximum, the minimum, and the range for each agency. Discuss the performances of these agencies. Which agency is the best performer? Which is the worst? What makes you draw these conclusions? If the state decides to retrain agencies one at a time, what would be your priority list of retraining?
2. Use the mean and standard deviation of the all cases to calculate z-scores for all test results.

Table 5–12

Number of Errors in Vehicle License Offices (per 100 licenses issued)

	Office A	Office B	Office C	Office D
Test 1	3.00	6.00	9.00	9.00
Test 2	7.00	3.00	4.00	9.00
Test 3	8.00	2.00	2.00	7.00
Test 4	5.00	5.00	2.00	6.00
Test 5	5.00	2.00	6.00	4.00
Test 6	5.00	9.00	2.00	1.00
Test 7	1.00	6.00	10.00	2.00
Test 8	9.00	4.00	5.00	9.00
Test 9	10.00	8.00	6.00	10.00
Test 10	5.00	2.00	7.00	4.00
Test 11	3.00	2.00	7.00	1.00
Test 12	4.00	1.00	8.00	2.00

Table 5–13

Performance Monitoring of Police Outputs

Monitoring Goal: To improve police response to crime
Monitoring Subject: Crime-responding activities or outputs of police department
Monitoring Question: Has there been an unusual increase in police crime-responding activities recently?
Output Measure 1: Number of calls for services dispatched (911 calls and officer initiated calls included)
Reporting Date: November 1, this year

	Monthly this year	Monthly last year
January	2764	2261
February	2201	2130
March	2029	2458
April	2815	2361
May	3769	3740
June	3917	3928
July	3726	3462
August	3798	3221
September	3193	3520
October	2277	2345
November	NA	2871
December	NA	2347
Total	30,489	34,644

Practice Problem 5–2

An ideal performance monitoring system should track not only outcomes but also inputs, process, and outputs. In the police response example in this chapter, we use a case of outcome monitoring on the crime rate. Table 5–13 shows the data of an output measure of that department: the number of calls for services dispatched. (Note that the measure can be treated as a process measure as well.)

1. Conduct a performance monitoring of the output using the tools of monitoring against performance standards and monitoring performance variation.
2. Write a brief paragraph to discuss whether there is a warning sign of underperformance.

Practice Problem 5–3

Conduct an outcome performance monitoring in an agency of your choice.