Chapter 8

Nonlinear Programming and Evolutionary Optimization

Learning Objectives

After reading this chapter students should be able to:

1. Explain the differences between linear and nonlinear optimization problems.
2. Describe the difficulties that may arise in solving nonlinear optimization problems.
3. Describe how the generalized reduced gradient (GRG) algorithm works.
4. Differentiate between global and local optimal solutions.
5. Describe the difference between convex and nonconvex feasible regions.
6. Describe the difference between convex and concave functions.
7. Describe the difference between smooth and non-smooth optimization problems.
8. Formulate and solve a variety of nonlinear optimization problems.
9. Interpret the sensitivity report for the optimal solution to a nonlinear programming (NLP) problem.
10. Describe how evolutionary/genetic algorithms work.
11. Describe the purpose and use of the "alldifferent" constraint type.
12. Explain and use the following functions: AVERAGE( ), COUNTIF( ), COVAR( ), IFERROR( ), INDEX( ), MAX( ), MMULT( ), PRODUCT( ), PsiOptParam( ), SQRT( ), SUM( ), SUMIF( ), SUMPRODUCT( ), VAR( ), VLOOKUP( ).

8-0 Introduction

Up to this point in our study of optimization, we have considered only mathematical programming models in which the objective function and constraints are *linear* functions of the decision variables. In many decision problems, the use of such linear functions is appropriate. Other types of optimization problems involve objective functions and constraints that *cannot* be modeled adequately using linear functions of the decision variables. These types of problems are called **nonlinear programming** (NLP) problems.

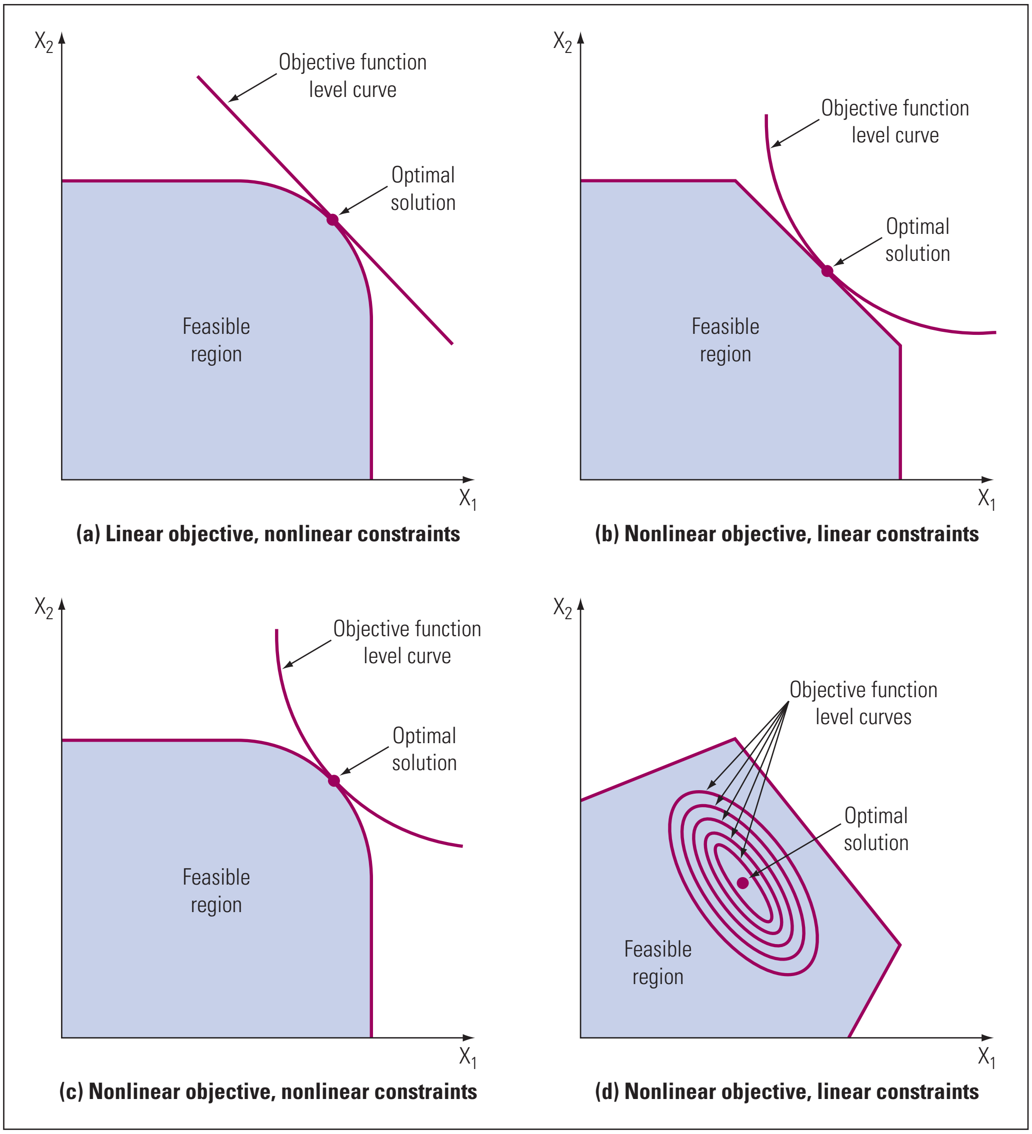
The process of formulating an NLP problem is virtually the same as formulating an LP problem. In each case, you must identify the appropriate decision variables and formulate an appropriate objective function and constraints using these variables. As you will see, the process of implementing and solving NLP problems in a spreadsheet is also similar to that for LP problems. However, the mechanics (that is, mathematical procedures) involved in solving NLP problems are very different. Although optimization software makes this difference somewhat transparent to the user of such systems, it is important to understand these differences so you can understand the difficulties you might encounter when solving an NLP problem. This chapter discusses some of the unique features and challenges involved in solving NLP problems, and presents several examples of managerial decision-making problems that can be modeled as NLP problems.

8-1 The Nature of NLP Problems

The main difference between an LP and NLP problem is that NLPs can have a nonlinear objective function and/or one or more nonlinear constraints. To understand the differences and difficulties nonlinearities introduce to an optimization problem, consider the various hypothetical NLP problems shown in Figure 8.1.

FIGURE 8.1

Examples of NLP problems with optimal solutions not at a corner point of the feasible region



The first graph in Figure 8.1, labeled (a), illustrates a problem with a linear objective function and a *nonlinear* feasible region. Note that the boundary lines of the feasible region for this problem are not all straight lines. At least one of the constraints in this problem must be nonlinear to cause the curve in the boundary line of the feasible region. This curve causes the unique optimal solution to this problem to occur at a solution that is not a corner point of the feasible region.

The second graph in Figure 8.1, labeled (b), shows a problem with a *nonlinear* objective function and a linear constraint set. As indicated in this graph, if an NLP problem has a nonlinear objective function, the level curves associated with the objective are also nonlinear. So from this graph, we observe that a nonlinear objective can cause the optimal solution to the NLP problem to occur at a solution that is not a corner point of the feasible region—even if all the constraints are linear.

The third graph in Figure 8.1, labeled (c), shows a problem with a *nonlinear* objective and a *nonlinear* constraint set. Here again, we see that the optimal solution to this NLP problem occurs at a solution that is not a corner point of the feasible region.

Finally, the fourth graph in Figure 8.1, labeled (d), shows another problem with a *nonlinear* objective and a *linear* constraint set. The optimal solution to this problem occurs at a point in the interior of the feasible region.

These graphs illustrate the major difference between LP and NLP problems—an optimal solution to an LP problem always occurs at a corner point of its feasible region, but this is not true of NLP problems. The optimal solution to some NLP problems might not occur on the boundary of the feasible region at all, but at some point in the interior of the feasible region. Therefore, the strategy of searching the corner points of the feasible region employed by the simplex method to solve LP problems will not work with NLP problems. We need another strategy to solve NLP problems.

8-2 Solution Strategies for NLP Problems

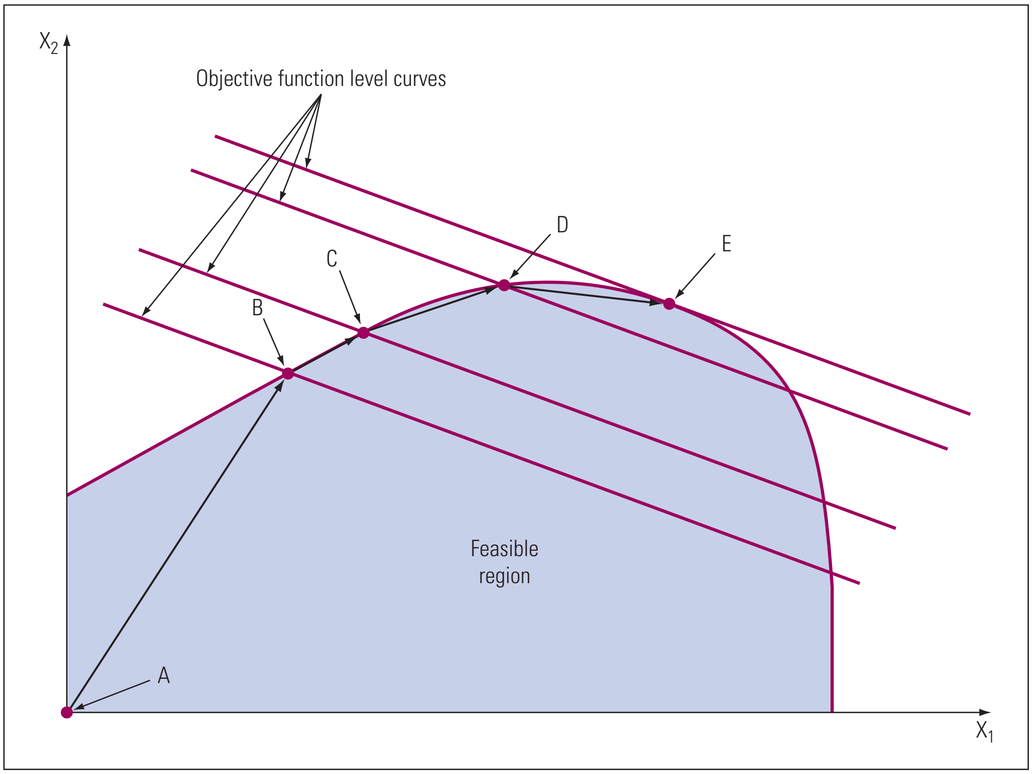
The solution procedure Solver uses to solve NLP problems is called the **generalized reduced gradient** (GRG) algorithm. The mathematics involved in this procedure is rather complex and goes beyond the scope and purpose of this text. However, the following discussion should give you a very basic (if somewhat imprecise) understanding of the ideas behind the GRG and other NLP solution algorithms.

NLP algorithms begin at any feasible solution to the NLP problem. This initial feasible solution is called the **starting point**. The algorithm then attempts to move from the starting point in a direction through the feasible region that causes the objective function value to improve. Some amount of movement (or a **step size**) in the selected feasible direction is then taken resulting in a new, and better, feasible solution to the problem. The algorithm next attempts to identify another feasible direction in which to move to obtain further improvements in the objective function value. If such a direction exists, the algorithm determines a new step size and moves in that direction to a new and better feasible solution. This process continues until the algorithm reaches a point at which there is no feasible direction in which to move that results in an improvement in the objective function. When no further possibility for improvement exists (or the potential for further improvement becomes arbitrarily small), the algorithm terminates.

Figure 8.2 shows a graphical example of how a crude NLP algorithm might work. In this graph, an initial feasible solution occurs at the origin (point A). The fastest rate of improvement in the objective function value occurs by moving from point A in the direction that is perpendicular to (or forms a 90-degree angle with) the level curves of the objective function. Feasible movement in this direction is possible from point A to point B where a boundary of the feasible region is encountered. From point B, moving along the edge of the feasible region to point C further improves the objective function value. At point C, the boundary of the feasible region begins to curve; therefore, continued movement in the direction from point B to point C is no longer feasible. From point C, a new direction through the interior of the feasible region allows movement to point D. This process continues from point D until the solution becomes arbitrarily close (or converges) to point E—the optimal solution.

FIGURE 8.2

Example of an NLP solution strategy



In moving from point A in Figure 8.2, we selected the direction that resulted in the fastest rate of improvement in the objective function. In retrospect, we can see that it would have been better to move from point A in the direction of point E. This direction does not result in the fastest rate of improvement in the objective as we move from point A, but it would have taken us to the optimal solution in a more direct fashion. Thus, it is not always best to move in the direction producing the fastest rate of improvement in the objective, nor is it always best to move as far as possible in that direction. The GRG algorithm used by Solver takes these issues into consideration as it determines the direction and step size of the movements to make. Thus, although the GRG algorithm usually cannot move directly from a starting point to an optimal solution, it does choose the path it takes in a more refined manner than outlined in Figure 8.2.

8-3 Local vs. Global Optimal Solutions

NLP solution algorithms terminate whenever they detect that no feasible direction exists in which it can move to produce a better objective function value (or when the amount of potential improvement becomes arbitrarily small). In such a situation, the current solution is a **local optimal solution**—a solution that is better than any other feasible solution in its immediate, or local, vicinity. However, a given local optimal solution might not be the best possible, or **global optimal**, solution to a problem. Another local optimal solution in some other area of the feasible region could be the best possible solution to the problem. This type of anomaly is illustrated in Figure 8.3.

If an NLP algorithm starts at point A in Figure 8.3, it could move immediately to point B and then along the feasible direction from point B to C. Because no feasible point in the vicinity of point C produces a better objective function value, point C is a local optimal solution and the algorithm terminates at this point. However, this is clearly *not* the best possible solution to

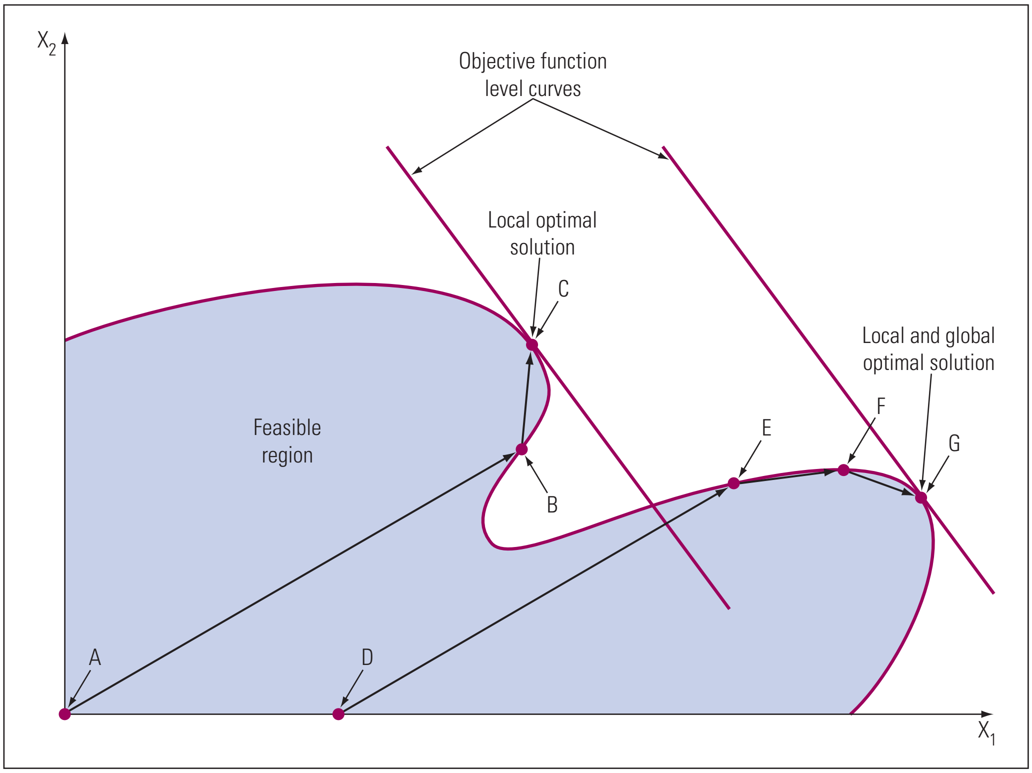


FIGURE 8.3

Local vs. global optimal solutions

this problem. If an NLP algorithm starts at point D in Figure 8.3, it could move immediately to point E, and then follow the feasible direction from point E to F and from point F to G. Note that point G is both a local and global optimal solution to this problem.

It is important to note that the feasible region of the problem in Figure 8.3 is nonconvex while those in Figures 8.1 and 8.2 are convex. A set of points X is called a convex set if for ***any*** two points in the set a straight line drawn connecting the two points falls entirely within X. (A line connecting points B and E in Figure 8.3 would not fall within the feasible region. Therefore the feasible region in Figure 8.3 is non-convex.) A function is convex (or concave) if the line connecting any two points on the function falls entirely above (or below) the function. Optimization problems with convex feasible regions and convex (or concave) objective functions are considerably easier to solve to global optimality than those that do not exhibit these properties.

Figure 8.3 highlights two important points about the GRG and all other NLP algorithms:

- NLP algorithms can terminate at a local optimal solution that might not be the global optimal solution to the problem.
- The local optimal solution at which an NLP algorithm terminates depends on the initial starting point.

The possibility of terminating at a local optimal solution is undesirable—but we have encountered this type of difficulty before. In our study of integer programming, we noted that suboptimal solutions to ILPs might be acceptable if they are within some allowable tolerance of the global optimal solution. Unfortunately, with NLP problems, it is difficult to determine how much worse a given local optimal solution is than the global optimal solution because most optimization packages do not provide a way of obtaining bounds on the optimal objective function values for these problems. However, many NLP problems have a single local optimal solution that, by definition, must also be the global optimal solution. So in many problems NLP algorithms will locate the global optimal solution but, as a general rule, we will not know whether the solution obtained is a global optimal solution. However, in Analytic Solver Platform some information about this issue can be obtained by running the convexity tester (by choosing Optimize > Analyze Original Model, or by clicking the "X-Checkmark" icon on the Model tab in the task pane). The result of convexity testing may be "Proven convex," "Proven non-convex," or "Nothing proven." If you see "Model Type - NLP Convex" in the Model Diagnosis area of the Task Pane Model tab, then you know that a local optimal solution is also a global optimal solution. If you see "NLP NonCvx" or just "NLP," then you have to assume that you have only a local optimal solution. In the non-convex case, it is usually a good idea to try starting NLP algorithms from different points to determine if the problem has different local optimal solutions. This procedure often reveals the global optimal solution. (Two questions at the end of this chapter illustrate this process.)

A Note about "Optimal" Solutions

When solving an NLP problem, Solver normally stops when the first of three numerical tests is satisfied, causing one of the following three completion messages to appear:

1. "**Solver found a solution. All constraints and optimality conditions are satisfied.**" This means Solver found a local optimal solution, but does not guarantee that the solution is the global optimal solution. Unless you know that a problem has only one local optimal solution (which must also be the global optimal solution), you should run Solver from several different starting points to increase the chances that you find the global optimal solution to your problem. The easiest way to do this is to set the Engine tab Global Optimization group MultiStart option to True before you solve—this will automatically run the Solver from several randomly (but efficiently) chosen starting points.
2. "**Solver has converged to the current solution. All constraints are satisfied.**" This means the objective function value changed very slowly for the last few iterations. If you suspect the solution is not a local optimal solution, your problem may be poorly scaled. The convergence option on the Engine tab can be reduced to avoid convergence at suboptimal solutions.
3. "**Solver cannot improve the current solution. All constraints are satisfied.**" This rare message means that your model is degenerate and the Solver is cycling. Degeneracy can often be eliminated by removing redundant constraints in a model.

A Note about Starting Points

Solver sometimes has trouble solving an NLP problem if it starts at the null starting point, where all the decision variables are set equal to 0—even if this solution is feasible. Therefore, when solving an NLP problem, it is best to specify a non-null starting solution whenever possible.

We will now consider several examples of NLP problems. These examples illustrate some of the differences between LP and NLP problems and provide insight into the broad range of problems that cannot be modeled adequately using LP.

8-4 Economic Order Quantity Models

The economic order quantity (EOQ) problem is one of the most common business problems for which nonlinear optimization can be used. This problem is encountered when a manager must determine the optimal number of units of a product to purchase whenever an order is placed. The basic model for an EOQ problem makes the following assumptions:

1. Demand for (or use of) the product is fairly constant throughout the year.
2. Each new order is delivered in full when the inventory level reaches 0.

Figure 8.4 illustrates the type of inventory patterns observed for a product when the preceding conditions are met. In each graph, the inventory levels are depleted at a constant rate, representing constant demand. Also, the inventory levels are replenished instantly whenever the inventory levels reach 0.

The key issue in an EOQ problem is to determine the optimal quantity to order whenever an order is placed for an item. The trade-offs in this decision are evident in Figure 8.4. The graphs

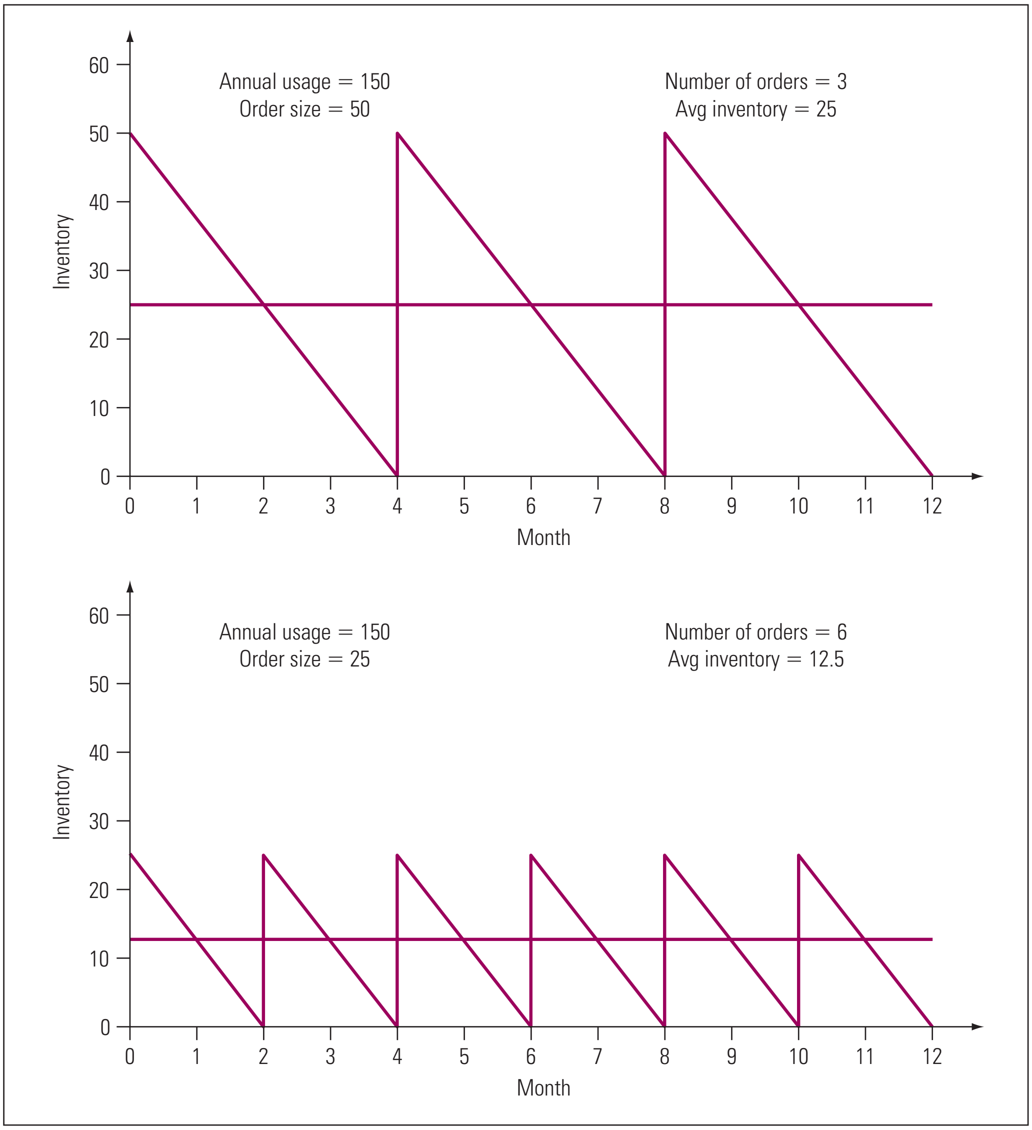
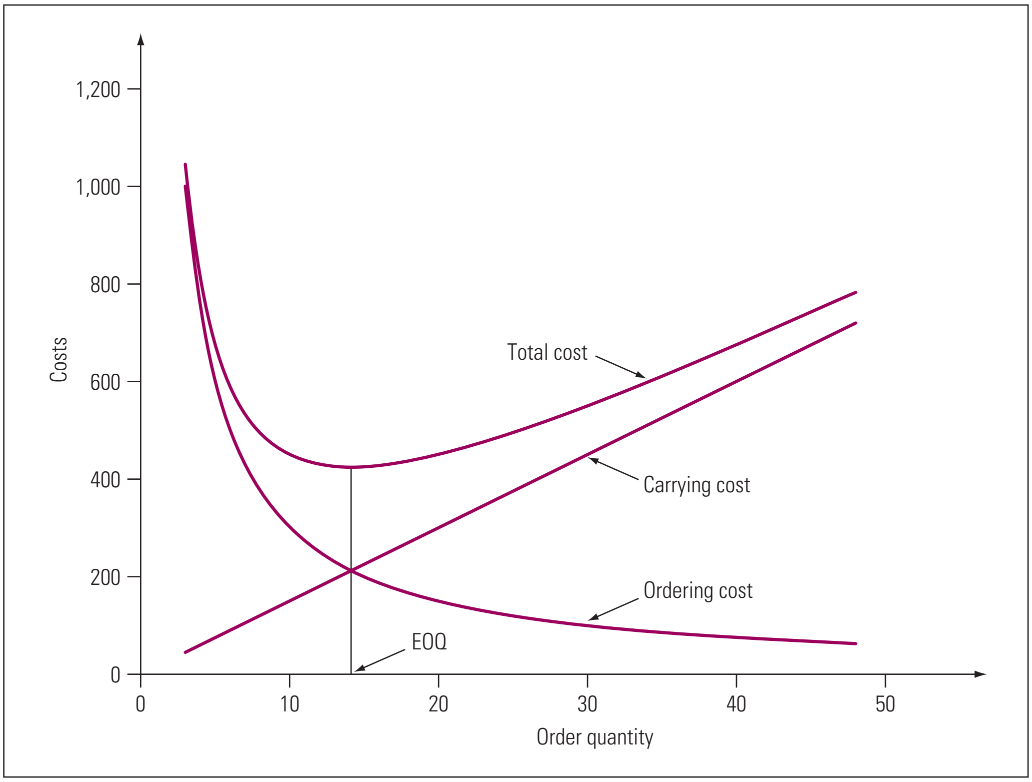


FIGURE 8.4

Inventory profiles of products for which the EOQ assumptions are met indicate two ways of obtaining 150 units of a product during the year. In the first graph, an order for 50 units is received whenever the inventory level drops to 0. This requires that three purchase orders be issued during the year and results in an average inventory level of 25 units. In the second graph, an order for 25 units is received whenever the inventory level drops to 0. This requires that six purchase orders be issued throughout the year and results in an average inventory level of 12.5 units. Thus, the first ordering strategy results in fewer purchase orders (and lower ordering costs) but higher inventory levels (and higher carrying costs). The second ordering strategy results in more purchase orders (and higher ordering costs) but lower levels of inventory (and lower carrying costs).

FIGURE 8.5

Relationship between order quantity, carrying cost, ordering cost, and total cost



In the basic EOQ model, the total annual cost of stocking a product is computed as the sum of the actual purchase cost of the product, plus the fixed cost of placing orders, plus the cost of holding (or carrying) the product in inventory. Figure 8.5 shows the relationships among order quantity, carrying cost, ordering cost, and total cost. Notice that as the order quantity increases, ordering costs decrease and carrying costs increase. The goal in this type of problem is to find the EOQ that minimizes the total cost.

The total annual cost of acquiring products that meet the stated assumptions is represented by:

$$\text{Total annual cost} = DC + \frac{D}{Q}S + \frac{Q}{2}Ci$$

where

D = annual demand for the item
C = unit purchase cost for the item
S = fixed cost of placing an order
i = cost of holding one unit in inventory for a year (expressed as a percentage of C)
Q = order quantity, or quantity ordered each time an order is placed

The first term in this formula (DC) represents the cost of purchasing a year's worth of the product. The second term $\frac{D}{Q}S$ represents the annual ordering costs. Specifically, $\frac{D}{Q}$ represents the number of orders placed during a year. Multiplying this quantity by S represents the cost of placing these orders. The third term $\frac{Q}{2}Ci$ represents the annual cost of holding inventory. On average, $\frac{Q}{2}$ units are held in inventory throughout the year (refer to Figure 8.4). Multiplying this term by Ci represents the cost of holding these units. The following example illustrates the use of the EOQ model.

Alan Wang is responsible for purchasing the paper used in all the copy machines and laser printers at the corporate headquarters of MetroBank. Alan projects that in the coming year he will need to purchase a total of 24,000 boxes of paper, which will be used at a fairly steady rate throughout the year. Each box of paper costs $35. Alan estimates that it costs $50 each time an order is placed (this includes the cost of placing the order plus the related costs in shipping and receiving). MetroBank assigns a cost of 18% to funds allocated to supplies and inventories because such funds are the lifeline of the bank and could be lent out to credit card customers who are willing to pay this rate on money borrowed from the bank. Alan has been placing paper orders once a quarter, but he wants to determine if another ordering pattern would be better. He wants to determine the most economical order quantity to use in purchasing the paper.

8-4a IMPLEMENTING THE MODEL

To solve this problem, we first need to create a spreadsheet model of the total cost formula described earlier, substituting the data for Alan's problem for the parameters D, C, S, and *i*. This spreadsheet implementation is shown in Figure 8.6 (and in the file Fig8-6.xlsm that accompanies this book).

In Figure 8.6, cell D4 represents the annual demand (D), cell D5 represents the per-unit cost (C), cell D6 represents the cost of placing an order (S), cell D7 represents the inventory holding cost (*i*) expressed as a percentage of an item's value, and cell D9 represents the order quantity (Q). The data corresponding to Alan's decision problem have been entered into the appropriate cells in this model. Because Alan places orders once a quarter (or four times a year), the order quantity in cell D9 is set at 24,000 ÷ 4 = 6,000.

We calculate each of the three pieces of our total cost function in cells D11, D12, and D13. Cell D11 contains the cost of purchasing a year's worth of paper, cell D12 represents the cost associated with placing orders, and cell D13 is the inventory holding cost that would be incurred. The sum of these costs is calculated in cell D14.

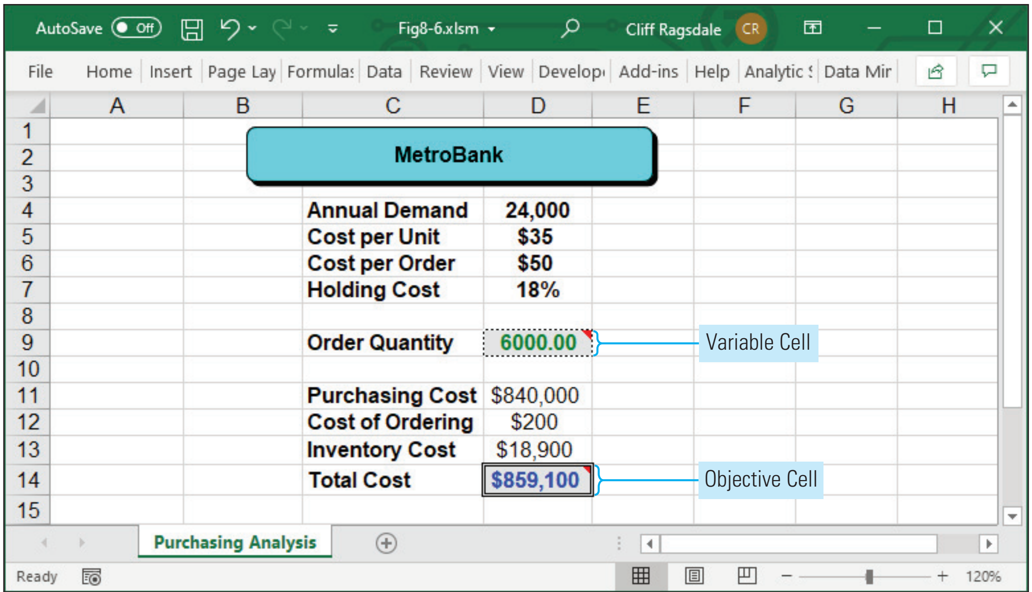
Formula for cell D11: = D5*D4
Formula for cell D12: = D4/D9*D6
Formula for cell D13: = D9/2*D7*D5
Formula for cell D14: = SUM(D11:D13)

8-4b SOLVING THE MODEL

The goal in this problem is to determine the order quantity (the value of Q) that minimizes the total cost. That is, we want Solver to determine the value for cell D9 that minimizes the value in cell D14. Figure 8.7 shows the Solver parameters and options required to solve this problem. Note that a constraint has been placed on cell D9 to prevent the order quantity from becoming 0 or negative. This constraint requires that at least one order must be placed during the year.

FIGURE 8.6

Spreadsheet implementation of MetroBank's paper purchasing problem



Key Cell Formulas

Cell	Formula	Copied to
D11	=D5*D4	--
D12	=D4/D9*D6	--
D13	=D9/2*D5*D7	--
D14	=SUM(D11:D13)	--

FIGURE 8.7

Solver parameters for MetroBank's paper purchasing problem

Solver Settings:

Objective: D14 (Min)
Variable cells: D9
Constraints:
D9 >=1

Solver Options:
Standard GRG Nonlinear Engine

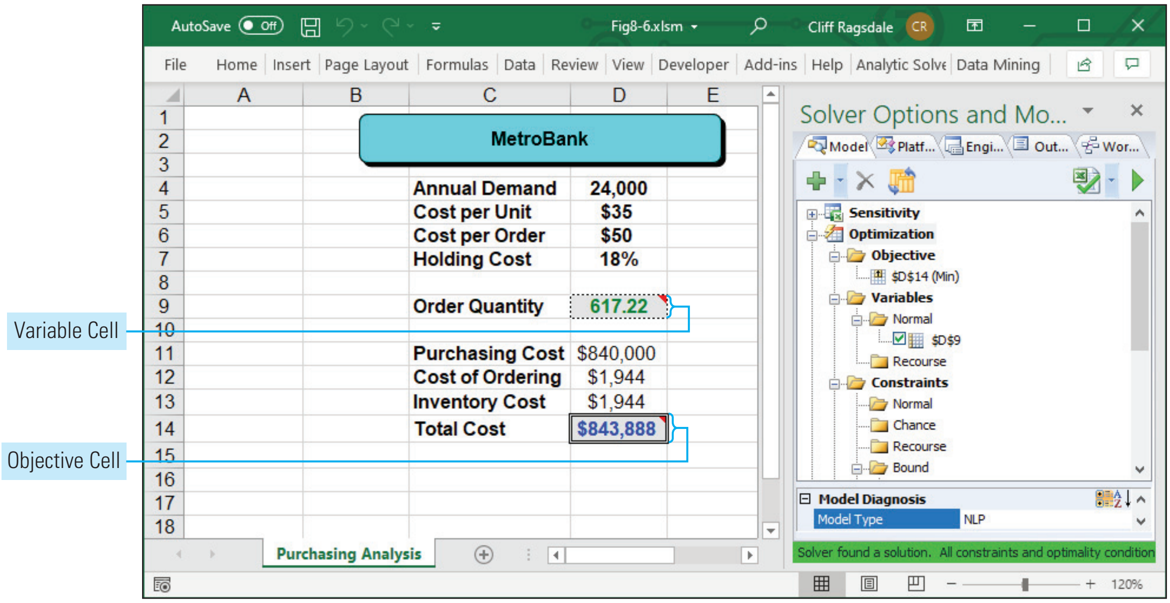
A Note about Engine Options

When solving an NLP problem, it is important *not* to select the Standard LP engine option. When this option is selected, Analytic Solver Platform conducts a number of internal tests to verify that the model is truly linear in the objective and constraints. If this option is selected and Solver's tests indicate that the model is *not* linear, a message appears indicating that the conditions for linearity are not satisfied.

8-4c ANALYZING THE SOLUTION

The optimal solution to this problem is shown in Figure 8.8. This solution indicates that the optimal number of boxes for Alan to order at any time is approximately 617. Because the total cost curve in the basic EOQ model has one minimum point, we can be sure that this local optimal solution is also the global optimal solution for this problem. Notice this solution occurs where the total ordering costs are in balance with the total holding costs. Using this order quantity, costs are reduced by approximately $15,211 from the earlier level shown in Figure 8.7 when an order quantity of 6,000 was used.

FIGURE 8.8 *Optimal solution to MetroBank's paper purchasing problem*



If Alan orders 617 boxes, he needs to place approximately 39 orders during the year (24,000 ÷ 617 = 38.89), or 1.333 orders per week (52 ÷ 39 = 1.333). As a practical matter, it might be easier for Alan to arrange for weekly deliveries of approximately 461 boxes. This would increase the total cost by only $167 to $844,055 but probably would be easier to manage and still save the bank more than $15,000 per year.

8-4d COMMENTS ON THE EOQ MODEL

There is another way to determine the optimal order quantity using the simple EOQ model. Using calculus, it can be shown that the optimal value of Q is represented by:

$$Q^* = \sqrt{\frac{2DS}{Ci}}$$

If we apply this formula to our example problem, we obtain:

$$Q^* = \sqrt{\frac{2 \times 24{,}000 \times 50}{35 \times 0.18}} = \sqrt{\frac{2{,}400{,}000}{6.3}} = 617.214$$

The value obtained using calculus is almost the same value obtained using Solver (refer to cell D9 in Figure 8.8). The slight difference in the results might be due to rounding, or to Solver stopping just short of converging on the exact solution.

Although the previous EOQ formula has its uses, we often must impose financial or storage space restrictions when determining optimal order quantities. The previous formula does not explicitly allow for such restrictions, but it is easy to impose these types of restrictions using Solver. In some of the problems at the end of this chapter, we will consider how the EOQ model can be adjusted to accommodate these types of restrictions, as well as quantity discounts. A complete discussion of the proper use and role of EOQ models in inventory control is beyond the scope of this text, but can be found in other texts devoted to production and operations management.

8-5 Location Problems

A number of decision problems involve determining the location of facilities or service centers. Examples might include determining the optimal location of manufacturing plants, warehouses, fire stations, or ambulance centers. The objective in these types of problems is often to determine a location that minimizes the distance between two or more service points. You might recall from basic algebra that the straight line (or Euclidean) distance between two points (X_1, Y_1) and (X_2, Y_2) on a standard X-Y graph is defined as:

$$\text{Distance} = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}$$

This type of calculation is likely to be involved in any problem in which the decision variables represent possible locations. The distance measure might occur in the objective function (e.g., we might want to minimize the distance between two or more points) or it might occur in a constraint (e.g., we might want to ensure that some minimum distance exists between two or more locations). Problems involving this type of distance measure are nonlinear. The following example illustrates the use of distance measures in a location problem.

> The Rappaport Communications Company provides cellular telephone services in several mid-western states. The company is planning to expand its customer base by offering cellular service in northeastern Ohio to the cities of Cleveland, Akron, Canton, and Youngstown. The company will install the hardware necessary to service customers in each city on preexisting communications towers in each city. The locations of these towers are summarized in Figure 8.9.
>
> However, the company also needs to construct a new communications tower somewhere between these cities to handle intercity calls. This tower will also allow cellular calls to be routed onto the satellite system for worldwide calling service. The tower the company is planning to build can cover areas within a 40-mile radius. Thus, the tower needs to be located within 40 miles of each of these cities.

It is important to note that we could have overlaid the X- and Y-axes on the map in Figure 8.9 in more than one way. The origin in Figure 8.9 could be located anywhere on the map without affecting the analysis. To establish the X-Y coordinates, we need an absolute reference point for the origin, but virtually any point on the map could be selected as the origin. Also, we can express the scaling of the X-axis and Y-axis in a number of ways: meters, miles, inches, feet, and so on. For our purposes, we will assume that each unit along the X- and Y-axes represents 1 mile.

8-5a DEFINING THE DECISION VARIABLES

In Figure 8.9, definite X-Y coordinates have been established to describe the locations of the cities. These points are fixed and are not under the decision maker's control. However, the

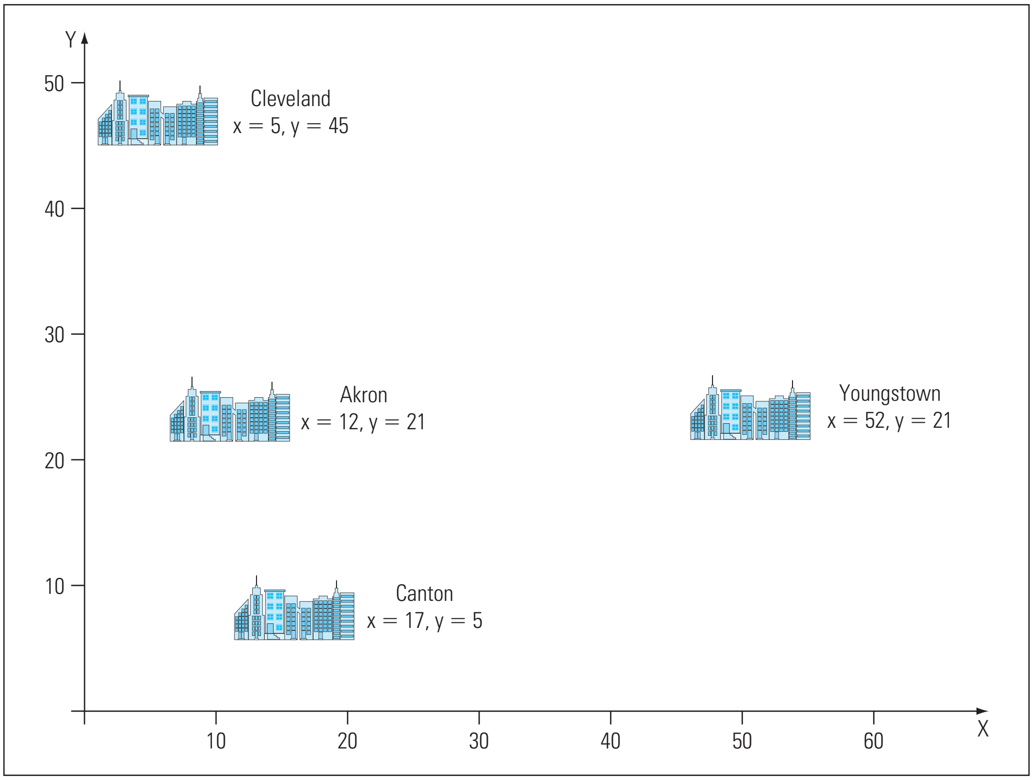


FIGURE 8.9

Map of Rappaport Communication's tower location problem

coordinates of the new communications tower have not been established. We will assume that Rappaport wants to determine the tower location that minimizes the total distance between the new tower and those in each of the four cities. (Note that this is equivalent to minimizing the average distance as well.) Thus, the coordinates of the new tower represent the decision variables in this problem, which are defined as:

$$X_1 = \text{location of the new tower with respect to the X-axis}$$
$$Y_1 = \text{location of the new tower with respect to the Y-axis}$$

8-5b DEFINING THE OBJECTIVE

The objective in this problem is to minimize the total distance from the new tower to each of the existing towers, defined as:

$$\begin{aligned}\text{MIN: } & \sqrt{(5 - X_1)^2 + (45 - Y_1)^2} + \sqrt{(12 - X_1)^2 + (21 - Y_1)^2} \\ & + \sqrt{(17 - X_1)^2 + (5 - Y_1)^2} + \sqrt{(52 - X_1)^2 + (21 - Y_1)^2}\end{aligned}$$

The first term in the objective calculates the distance from the tower in Cleveland, at X-Y coordinates (5, 45), to the location of the new tower, whose location is defined by the values X_1 and Y_1. The remaining terms perform similar calculations for the towers in Akron, Canton, and Youngstown.

8-5c DEFINING THE CONSTRAINTS

The problem statement noted that the new tower has a 40-mile transmission radius and therefore must be located within 40 miles of each of the existing towers. The following constraints ensure that the distance from each of the existing towers to the new tower is no larger than 40 miles.

$$\sqrt{(5 - X_1)^2 + (45 - Y_1)^2} \le 40 \quad \} \text{ Cleveland distance constraint}$$
$$\sqrt{(12 - X_1)^2 + (21 - Y_1)^2} \le 40 \quad \} \text{ Akron distance constraint}$$
$$\sqrt{(17 - X_1)^2 + (5 - Y_1)^2} \le 40 \quad \} \text{ Canton distance constraint}$$
$$\sqrt{(52 - X_1)^2 + (21 - Y_1)^2} \le 40 \quad \} \text{ Youngstown distance constraint}$$

Graphically, these constraints would be drawn as four circles, each with a 40-mile radius, each centered at one of the of the four existing tower locations. The intersection of these circles would represent the feasible region for the problem.

8-5d IMPLEMENTING THE MODEL

In summary, the problem Rappaport Communications wants to solve is:

$$\text{MIN: } \sqrt{(5 - X_1)^2 + (45 - Y_1)^2} + \sqrt{(12 - X_1)^2 + (21 - Y_1)^2}$$
$$+ \sqrt{(17 - X_1)^2 + (5 - Y_1)^2} + \sqrt{(52 - X_1)^2 + (21 - Y_1)^2}$$

Subject to:

$$\sqrt{(5 - X_1)^2 + (45 - Y_1)^2} \le 40 \quad \} \text{ Cleveland distance constraint}$$
$$\sqrt{(12 - X_1)^2 + (21 - Y_1)^2} \le 40 \quad \} \text{ Akron distance constraint}$$
$$\sqrt{(17 - X_1)^2 + (5 - Y_1)^2} \le 40 \quad \} \text{ Canton distance constraint}$$
$$\sqrt{(52 - X_1)^2 + (21 - Y_1)^2} \le 40 \quad \} \text{ Youngstown distance constraint}$$

Note that both the objective and constraints for this problem are nonlinear. One approach to implementing the model for this problem in a spreadsheet is shown in Figure 8.10 (and in the file Fig8-10.xlsm that accompanies this book).

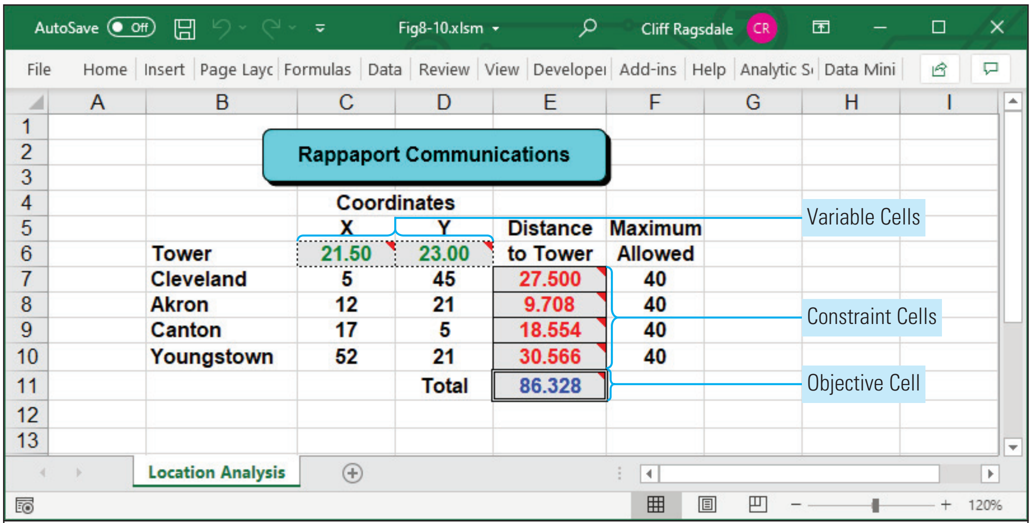
In this spreadsheet, cells C6 and D6 are used to represent the decision variables X_1 and Y_1, which correspond to the X-Y coordinates of the location of the new tower. The locations of the existing towers are listed in terms of their X-Y coordinates in rows 7 through 10 of columns C and D. Reasonable starting values for X_1 and Y_1 in this problem would be the average values of the X and Y coordinates of the existing tower locations. These averages were computed and entered in cells C6 and D6.

Column E calculates the distance from each existing tower to the selected location for the new tower. Specifically, cell E7 contains the following formula, which is copied to cells E8 through E10:

Formula for cell E7: =SQRT((C7 − $C$6)^2 + (D7 − $D$6)^2)
(Copy to E8 through E10.)

These cells represent the LHS formulas for the problem. The RHS values for these constraints are given in cells F7 through F10. The objective function for the problem is then implemented easily in cell E11 with the formula:

Formula for cell E11: =SUM(E7:E10)



Key Cell Formulas

Cell	Formula	Copied to
E7	=SQRT((C7−$C$6)^2+(D7−$D$6)^2)	E8:E10
E11	=SUM(E7:E10)	--

FIGURE 8.10

Spreadsheet implementation of the tower location problem

8-5e SOLVING THE MODEL AND ANALYZING THE SOLUTION

Figure 8.11 shows the Solver settings and options used to solve this problem, and Figure 8.12 shows the optimal solution.

Solver Settings:

Objective: E11 (Min)
Variable cells: C6:D6
Constraints:
  E7:E10 <= F7:F10

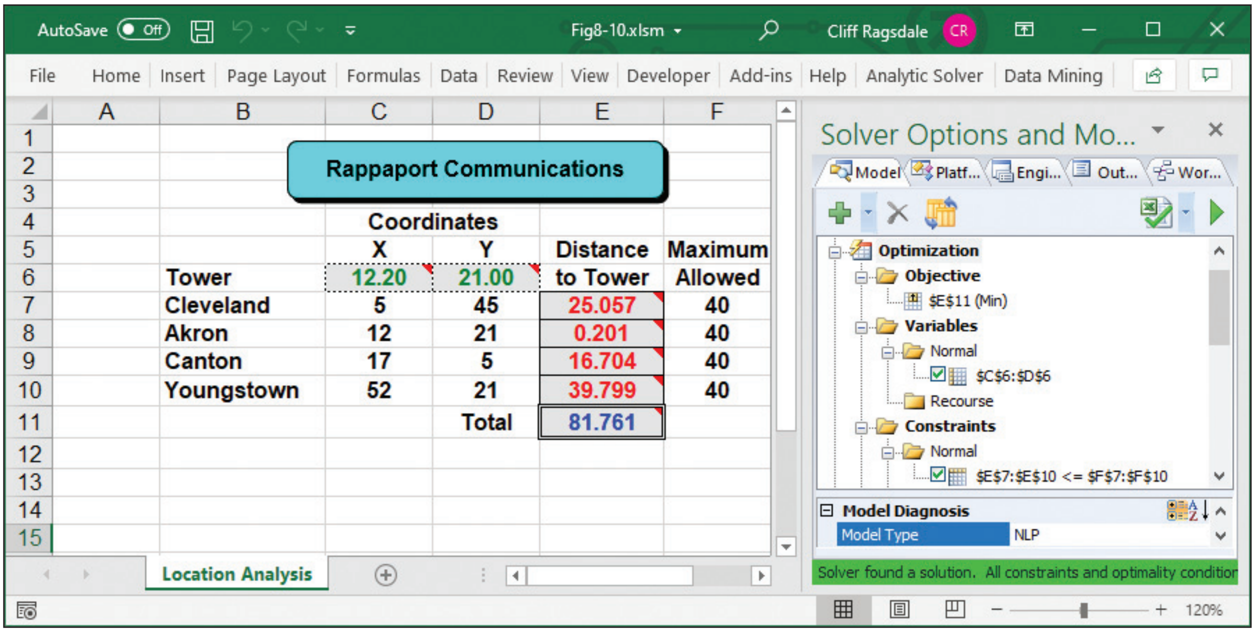
Solver Options:
  Standard GRG Nonlinear Engine

FIGURE 8.11

Solver parameters for the tower location problem

The solution in Figure 8.12 indicates that if the new tower is located at the coordinates $X_1 = 12.2$ and $Y_1 = 21.0$, the total distance between the towers is 81.761 miles (so the average distance is 20.4 miles). If you try re-solving this problem from a variety of starting points, you can verify that this is the global optimal solution to the problem. Interestingly, the coordinates of this location for the new tower are virtually identical to the coordinates of the existing tower in Akron. So, the solution to this problem may not involve building a new tower at all but, instead, Rappaport may want to investigate the feasibility of upgrading or retrofitting the existing Akron tower to play the role of the "new" tower.

FIGURE 8.12 *Optimal solution to the tower location problem*



8-5f ANOTHER SOLUTION TO THE PROBLEM

The solution shown in Figure 8.12 positions the new tower at essentially the same location as the existing tower in Akron. So, instead of building a new tower, perhaps the company should consider upgrading the tower at Akron with the equipment needed for handling intercity calls. On the other hand, the distance from Akron to the tower in Youngstown is almost 40 miles—the limit of the broadcast radius of the new equipment. Thus, Rappaport may prefer a solution that provides more of a safety margin on their broadcast range, to help ensure quality and reliability during inclement weather.

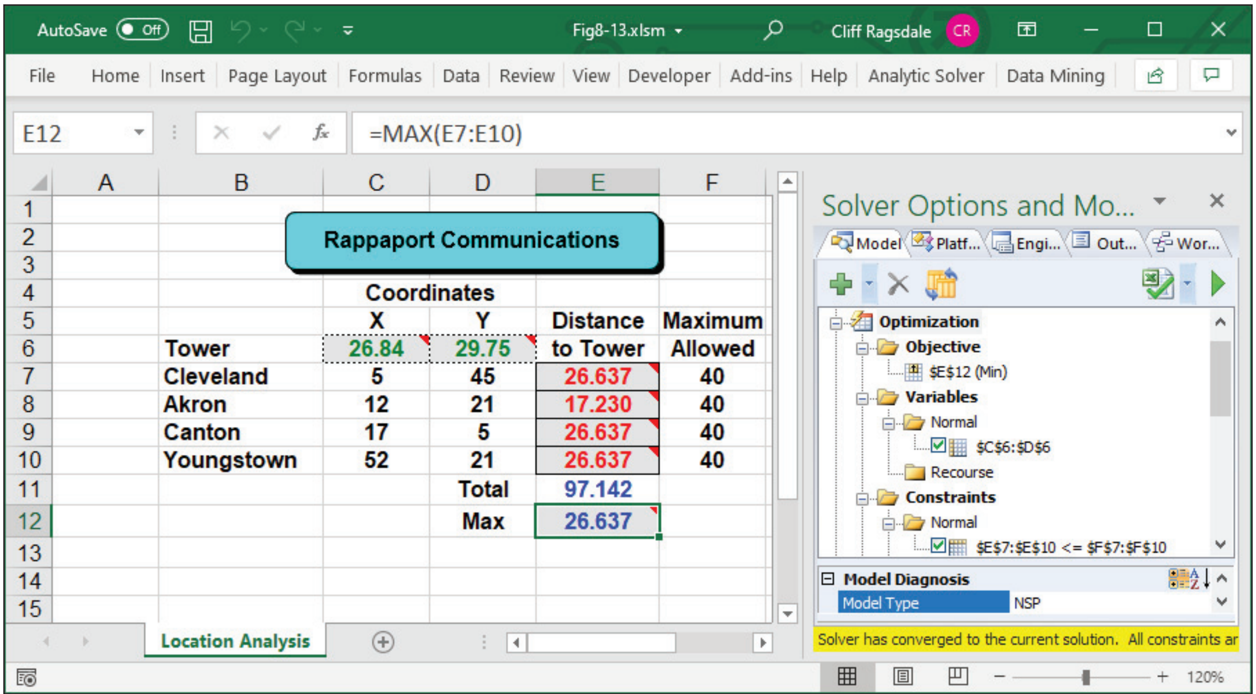
Another objective that could be applied to this problem would attempt to minimize the maximum distance from the new tower to each of the existing towers. Figure 8.13 (and the file Fig8-13.xlsm that accompanies this book) shows the solution to this new problem.

To obtain this solution, we implemented a new objective function in cell E12 to compute the maximum of the distances in column E as follows:

Formula for cell E12: =MAX(E7:E10)

We then instructed Solver to minimize E12 to obtain the solution shown. This solution positions the new tower at the X-Y coordinates (26.84, 29.75). Although the total distance associated with this solution increased to 97.142 (or an average of 24.28 miles), the maximum distance was reduced to about 26.6 miles. Thus, Rappaport might prefer this solution to the alternative shown in Figure 8.12.

FIGURE 8.13 *Another solution to the tower location problem minimizing the maximum distance*



Key Cell Formulas

Cell	Formula	Copied to
E7	=SQRT((C7−$C$6)^2+(D7−$D$6)^2)	E8:E10
E11	=SUM(E7:E10)	--
E12	=MAX(E7:E10)	

A Note about Non-Smooth Optimization Problems

In Figure 8.13, notice that the revised model containing the formula MAX(E7:E10) in cell E12 is diagnosed as an NSP—or a non-smooth optimization problem. Non-smooth optimization problems often contain EXCEL functions like IF( ), MAX( ), MIN( ), CHOOSE( ), or LOOKUP( ). Problems containing these (and similar) functions are non-smooth in the sense that their derivatives are not continuous. Generally speaking, with non-smooth problems you can have confidence that the solutions are "good" but they are not guaranteed to be globally or even locally optimal.

8-5g SOME COMMENTS ABOUT THE SOLUTION TO LOCATION PROBLEMS

When solving location problems, it is possible that the location indicated by the optimal solution may simply not work. For instance, the land at the optimal location might not be for sale, the "land" at this location might be a lake, the land might be zoned for residential purposes only, etc. However, solutions to location problems often give decision makers a good idea about where to start looking for suitable property to acquire for the problem at hand. It might also be possible to add constraints to location problems that eliminate certain areas from consideration if they are inappropriate or unavailable.

8-6 Nonlinear Network Flow Problem

In Chapter 5, we looked at several different types of network flow problems with linear objective functions and linear constraint sets. We noted that the constraints in network flow models have a special structure in which the flow into a node must be balanced with the flow out of the same node. Numerous decision problems exist in which the balance-of-flow restrictions must be maintained while optimizing a nonlinear objective function. We present one such example here.

> SafetyTrans is a trucking company that specializes in transporting extremely valuable and extremely hazardous materials. Due to the nature of its business, the company places great importance on maintaining a clean driving safety record. This not only helps keep their reputation up, but also helps keep their insurance premiums down. The company is also conscious of the fact that when carrying hazardous materials, the environmental consequences of even a minor accident could be disastrous.
>
> Whereas most trucking companies are interested in identifying routes that provide for the quickest or least costly transportation, SafetyTrans likes to ensure that it selects routes that are the least likely to result in an accident. The company is currently trying to identify the safest routes for carrying a load of hazardous materials from Los Angeles, California to Amarillo, Texas. The network in Figure 8.14 summarizes the routes under consideration. The numbers on each arc represent the probability of having an accident on each potential leg of the journey. SafetyTrans maintains a national database of such probabilities developed from data they receive from the National Highway Safety Administration and the various Departments of Transportation in each state.

8-6a DEFINING THE DECISION VARIABLES

The problem summarized in Figure 8.14 is very similar to the shortest path problem described in Chapter 5. As in the shortest path problem, here we will need one variable for each of the arcs (or routes) in the problem. Each decision variable will indicate whether or not a particular route is used. We will define these variables as follows:

$$Y_{ij} = \begin{cases} 1, \text{ if the route from node } i \text{ to node } j \text{ is selected} \\ 0, \text{ otherwise} \end{cases}$$

8-6b DEFINING THE OBJECTIVE

The objective in this problem is to find the route that minimizes the probability of having an accident, or equivalently, the route that maximizes the probability of not having an accident. Let

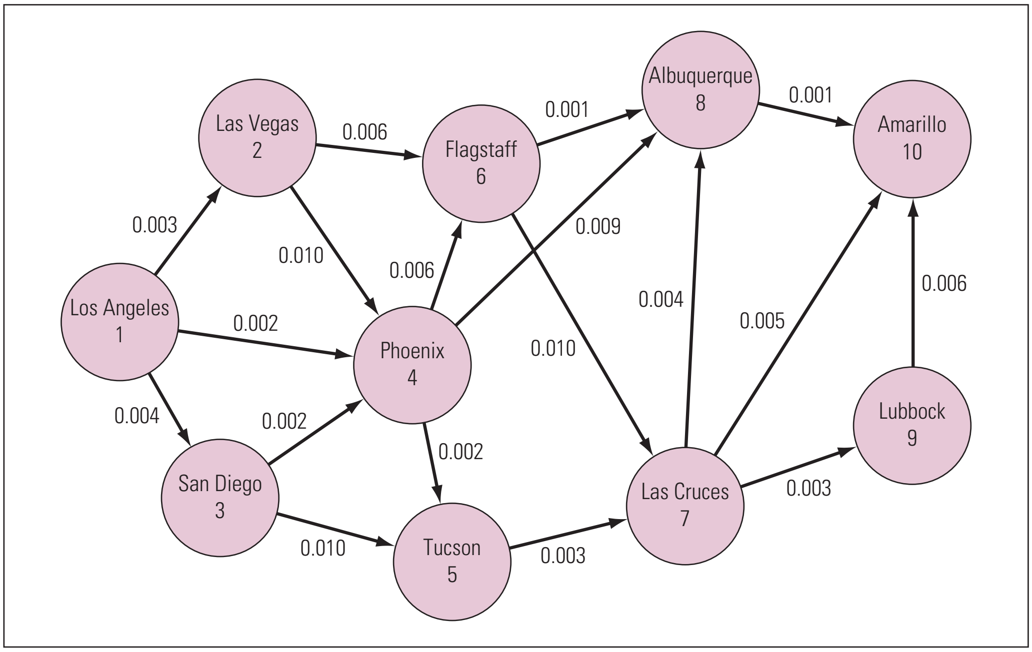


FIGURE 8.14

Network representation of the SafetyTrans routing problem

P_{ij} = the probability of having an accident while traveling from node *i* to node *j*. Then, the probability of ***not*** having an accident while traveling from node *i* to node *j* is $1-P_{ij}$. For example, the probability of not having an accident while traveling from Los Angeles to Las Vegas is $1 - P_{12} = 1 - 0.003 = 0.997$. The objective of maximizing the probability of not having an accident is given by:

$$\text{MAX: } (1 - P_{12}Y_{12})\,(1 - P_{13}Y_{13})\,(1 - P_{14}Y_{14})\,(1 - P_{24}Y_{24})\,(1 - P_{26}Y_{26}) \cdots (1 - P_{9,10}Y_{9,10})$$

The first term in this objective returns the value 1 if $Y_{12} = 0$ and the value $1 - P_{12}$ if $Y_{12} = 1$. Thus, if we take the route from Los Angeles to Las Vegas ($Y_{12} = 1$), the value 0.997 is multiplied by the remaining terms in the objective function. If we do not take the route from Los Angeles to Las Vegas ($Y_{12} = 0$), the value 1 is multiplied by the remaining terms in the objective function. (Of course, multiplying by 1 has no effect.) The remaining terms in the objective have similar interpretations. So, this objective function computes the probabilities of not having accidents on whichever routes are selected and then computes the products of these values. The result is the overall probability of not having an accident on any set of selected routes. This is the value SafetyTrans wants to maximize.

8-6c DEFINING THE CONSTRAINTS

To solve a shortest path network flow problem, we assign the starting node a supply value of −1, assign the ending node a demand value of +1, and apply the balance-of-flow rule covered in Chapter 5. This results in the following set of constraints for our example problem.

$$-Y_{12} - Y_{13} - Y_{14} = -1 \quad \}\text{ balance-of-flow constraint for node 1}$$

$$+Y_{12} - Y_{24} - Y_{26} = 0 \quad \}\text{ balance-of-flow constraint for node 2}$$

$$+Y_{13} - Y_{34} - Y_{35} = 0 \quad \}\text{ balance-of-flow constraint for node 3}$$

$$
\begin{array}{lcll}
+Y_{14} + Y_{24} + Y_{34} - Y_{45} - Y_{46} - Y_{48} & = & 0 & \} \text{ balance-of-flow constraint for node 4} \\
+Y_{35} + Y_{45} - Y_{57} & = & 0 & \} \text{ balance-of-flow constraint for node 5} \\
+Y_{26} + Y_{46} - Y_{67} - Y_{68} & = & 0 & \} \text{ balance-of-flow constraint for node 6} \\
+Y_{57} + Y_{67} - Y_{78} - Y_{79} - Y_{7,10} & = & 0 & \} \text{ balance-of-flow constraint for node 7} \\
+Y_{48} + Y_{68} + Y_{78} - Y_{8,10} & = & 0 & \} \text{ balance-of-flow constraint for node 8} \\
+Y_{79} - Y_{9,10} & = & 0 & \} \text{ balance-of-flow constraint for node 9} \\
+Y_{7,10} + Y_{8,10} + Y_{9,10} & = & 1 & \} \text{ balance-of-flow constraint for node 10}
\end{array}
$$

The first constraint ensures that one unit flows out of node 1 to nodes 2, 3, or 4. The last constraint ensures that one unit flows into node 10 from nodes 7, 8, or 9. The remaining constraints ensure that any flow into nodes 2 through 9 is balanced by an equal amount of flow out of those nodes.

8-6d IMPLEMENTING THE MODEL

In summary, the problem SafetyTrans wants to solve is:

$$
\text{MAX:} \quad (1 - 0.003Y_{12})\,(1 - 0.004Y_{13})\,(1 - 0.002Y_{14})\,(1 - 0.010Y_{24})\,(1 - 0.006Y_{26}) \cdots (1-0.006Y_{9,10})
$$

Subject to:

$$
\begin{array}{lcll}
-Y_{12} - Y_{13} - Y_{14} & = & -1 & \} \text{ balance-of-flow constraint for node 1} \\
+Y_{12} - Y_{24} - Y_{26} & = & 0 & \} \text{ balance-of-flow constraint for node 2} \\
+Y_{13} - Y_{34} - Y_{35} & = & 0 & \} \text{ balance-of-flow constraint for node 3} \\
+Y_{14} + Y_{24} + Y_{34} - Y_{45} - Y_{46} - Y_{48} & = & 0 & \} \text{ balance-of-flow constraint for node 4} \\
+Y_{35} + Y_{45} - Y_{57} & = & 0 & \} \text{ balance-of-flow constraint for node 5} \\
+Y_{26} + Y_{46} - Y_{67} - Y_{68} & = & 0 & \} \text{ balance-of-flow constraint for node 6} \\
+Y_{57} + Y_{67} - Y_{78} - Y_{79} - Y_{7,10} & = & 0 & \} \text{ balance-of-flow constraint for node 7} \\
+Y_{48} + Y_{68} + Y_{78} - Y_{8,10} & = & 0 & \} \text{ balance-of-flow constraint for node 8} \\
+Y_{79} - Y_{9,10} & = & 0 & \} \text{ balance-of-flow constraint for node 9} \\
+Y_{7,10} + Y_{8,10} + Y_{9,10} & = & 1 & \} \text{ balance-of-flow constraint for node 10}
\end{array}
$$

All Y_{ij} binary

One approach to implementing this model is shown in Figure 8.15. In this spreadsheet, cells A6 through A23 represent our decision variables.

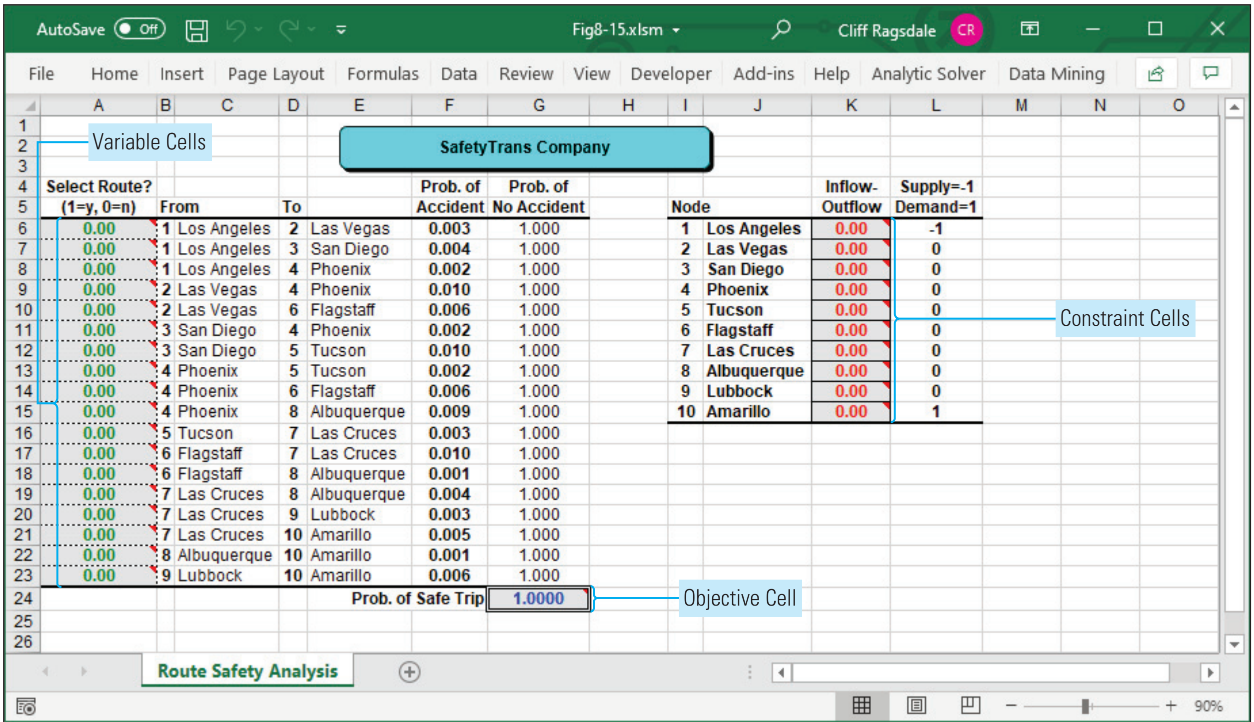
The LHS formulas for the constraints in this problem are implemented in cells K6 through K15 using the same technique as described in Chapter 5. The RHS for the constraints are given in cells L6 through L15. Specifically, we enter the following formula in cells K6 and copy down the rest of the column:

Formula for cell K6: =SUMIF($D$6:$D$23,I6,$A$6:$A$23)
(Copy to cells K7 through K15.) −SUMIF($B$6:$B$23,I6,$A$6:$A$23)

The probabilities of having an accident on each of the routes are listed in cells F6 through F23. Each of the terms for the objective function was then implemented in cells G6 through G23 as follows:

Formula for cell G6: = 1 − A6*F6
(Copy to cells G7 through G23.)

FIGURE 8.15 *Spreadsheet implementation of the SafetyTrans routing problem*



Select Route? (1=y, 0=n)	From		To		Prob. of Accident	Prob. of No Accident
0.00	1	Los Angeles	2	Las Vegas	0.003	1.000
0.00	1	Los Angeles	3	San Diego	0.004	1.000
0.00	1	Los Angeles	4	Phoenix	0.002	1.000
0.00	2	Las Vegas	4	Phoenix	0.010	1.000
0.00	2	Las Vegas	6	Flagstaff	0.006	1.000
0.00	3	San Diego	4	Phoenix	0.002	1.000
0.00	3	San Diego	5	Tucson	0.010	1.000
0.00	4	Phoenix	5	Tucson	0.002	1.000
0.00	4	Phoenix	6	Flagstaff	0.006	1.000
0.00	4	Phoenix	8	Albuquerque	0.009	1.000
0.00	5	Tucson	7	Las Cruces	0.003	1.000
0.00	6	Flagstaff	7	Las Cruces	0.010	1.000
0.00	6	Flagstaff	8	Albuquerque	0.001	1.000
0.00	7	Las Cruces	8	Albuquerque	0.004	1.000
0.00	7	Las Cruces	9	Lubbock	0.003	1.000
0.00	7	Las Cruces	10	Amarillo	0.005	1.000
0.00	8	Albuquerque	10	Amarillo	0.001	1.000
0.00	9	Lubbock	10	Amarillo	0.006	1.000
					Prob. of Safe Trip	1.0000

Node		Inflow-Outflow	Supply=-1 Demand=1
1	Los Angeles	0.00	-1
2	Las Vegas	0.00	0
3	San Diego	0.00	0
4	Phoenix	0.00	0
5	Tucson	0.00	0
6	Flagstaff	0.00	0
7	Las Cruces	0.00	0
8	Albuquerque	0.00	0
9	Lubbock	0.00	0
10	Amarillo	0.00	1

Key Cell Formulas

Cell	Formula	Copied to
C6	=VLOOKUP(B6,$I$6:$J$15,2)	C7:C23 and E6:E23
G6	=1−A6*F6	G7:G23
G24	=PRODUCT(G6:G23)	--
K6	=SUMIF($D$6:$D$23,I6,$A$6:$A$23) −SUMIF($B$6:$B$23,I6,$A$6:$A$23)	K7:K15

Note that the formula in G6 corresponds exactly to the first term in the objective $(1 - Y_{12}P_{12})$ as described earlier. Next, the product of the values in cells G6 through G23 is calculated in cell G24 as:

Formula for cell G24: =PRODUCT(G6:G23)

Figure 8.16 shows the Solver settings and options used to solve this problem. The optimal solution is shown in Figure 8.17.

8-6e SOLVING THE MODEL AND ANALYZING THE SOLUTION

The solution to this problem indicates $Y_{14} = Y_{46} = Y_{68} = Y_{8,10} = 1$, and all other $Y_{ij} = 0$. Thus, the optimal (safest) route is to travel from Los Angeles to Phoenix to Flagstaff to Albuquerque to Amarillo. Following this route, there is a 0.99 probability of not having an accident. Solving this

FIGURE 8.16

Solver parameters for the SafetyTrans problem

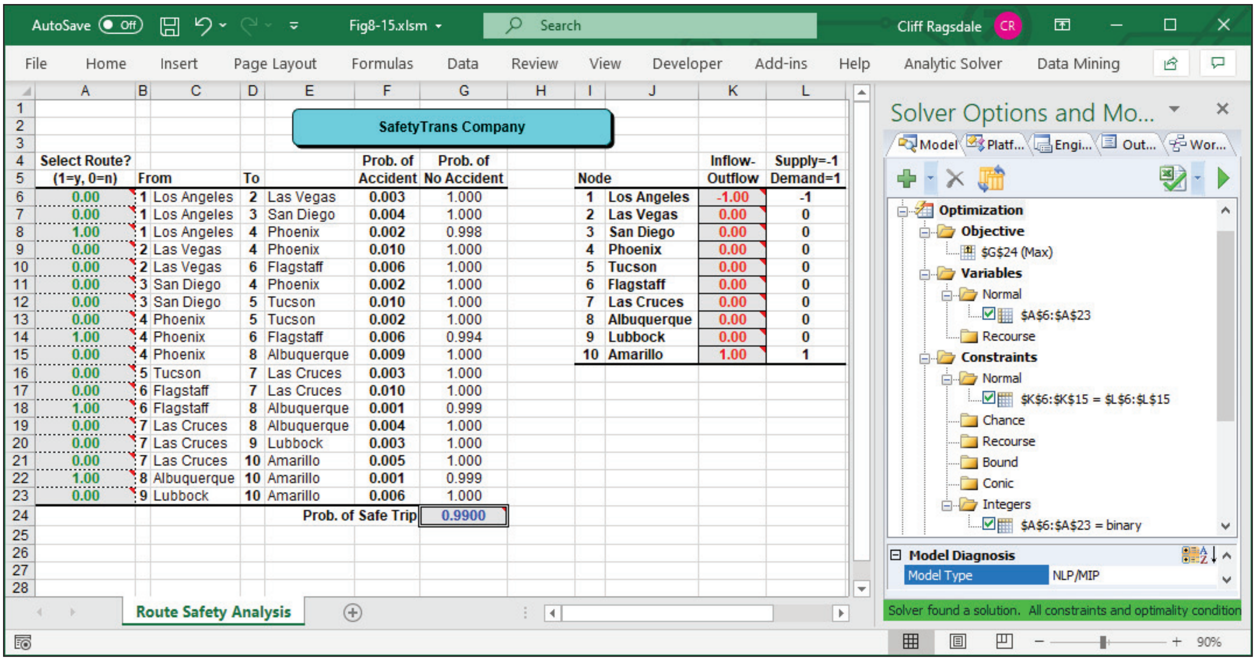
Solver Settings:

Objective: G24 (Max)
Variable cells: A6:A23
Constraints:
K6:K15 = L6:L15
A6:A23 = binary

Solver Options:

Standard GRG Nonlinear Engine
Integer Tolerance = 0

FIGURE 8.17 *Optimal solution to the SafetyTrans problem*



problem from numerous starting points indicates that this is the global optimal solution to the problem.

If you solve this model again minimizing the objective, you will discover that the least safe route has a 0.9626 probability of not having an accident. This may lead some to conclude that it doesn't make much difference which route is taken because the differences in the best case and worst case probabilities seem minimal. However, if it costs \$30,000,000 to clean up an accident involving hazardous waste, the expected cost of taking the safest route is $(1 - 0.99) \times \$30{,}000{,}000 = \$300{,}000$ and the expected cost of taking the least safe route is $(1 - 0.9626) \times \$30{,}000{,}000 = \$1{,}122{,}000$. So although the differences in probabilities may appear small, the differences in the potential outcomes can be quite significant. Of course, this doesn't even consider the potential loss of life and environmental damage that no amount of money can fix.

There are a number of other areas to which this type of nonlinear network flow model can be applied. Analysts are often interested in determining the "weakest link" in a telecommunications network or production system. The same type of problem described here could be solved to determine the least reliable path through these types of networks.

8-7 Project Selection Problems

In Chapter 6, we looked at a project selection example in which we wanted to select the combination of projects that produced the greatest net present value (NPV) subject to various resource restrictions. In these types of problems, there is often some uncertainty about whether a selected project can actually be completed successfully, and this success might be influenced by the amount of resources devoted to the project. The following example illustrates how NLP techniques can be used to help model this uncertainty in a selected project's potential for success.

The directors of the TMC Corporation are trying to determine how to allocate their R&D budget for the coming year. Six different projects are under consideration. The directors believe that the success of each project depends in part on the number of engineers assigned. Each project proposal includes an estimate of the probability of success as a function of the number of engineers assigned. Each probability function is of the form:

$$\text{P}_i = \text{probability of success for project } i \text{ if assigned } \text{X}_i \text{ engineers} = \frac{\text{X}_i}{\text{X}_i + \varepsilon_i}$$

where ε_i is a positive constant for project i that determines the shape of its probability function. The probability functions for several of the projects are shown in Figure 8.18. The following table summarizes the initial startup funds required for each project and the estimated NPV the project will generate if it is successful.

Project	**1**	**2**	**3**	**4**	**5**	**6**
Startup Costs	\$325	\$200	\$490	\$125	\$710	\$240
Net Present Value	\$750	\$120	\$900	\$400	\$1,100	\$800
Probability Parameter ε_i	3.1	2.5	4.5	5.6	8.2	8.5

(*Note:* All monetary values are in \$1000s.)

TMC's directors have agreed to hire up to 25 engineers to assign to these projects and are willing to allocate \$1.7 million of the R&D budget to cover the startup costs for selected projects. They want to determine the project selection and resource allocation strategy that will maximize the expected NPV.

8-7a DEFINING THE DECISION VARIABLES

TMC's directors must make two separate but related decisions. First, they must determine which projects to select. We will use the following binary variables to model these decisions:

$$\text{Y}_i = \begin{cases} 1, \text{ if project } i \text{ is selected} \\ 0, \text{ otherwise} \end{cases}, \quad i = 1, 2, 3, \ldots, 6$$

Second, the directors must determine the number of engineers to assign to each project. We will model these decisions with the following variables:

$$\text{X}_i = \text{the number of engineers to assign to project } i, \;\; i = 1, 2, 3, \ldots, 6$$

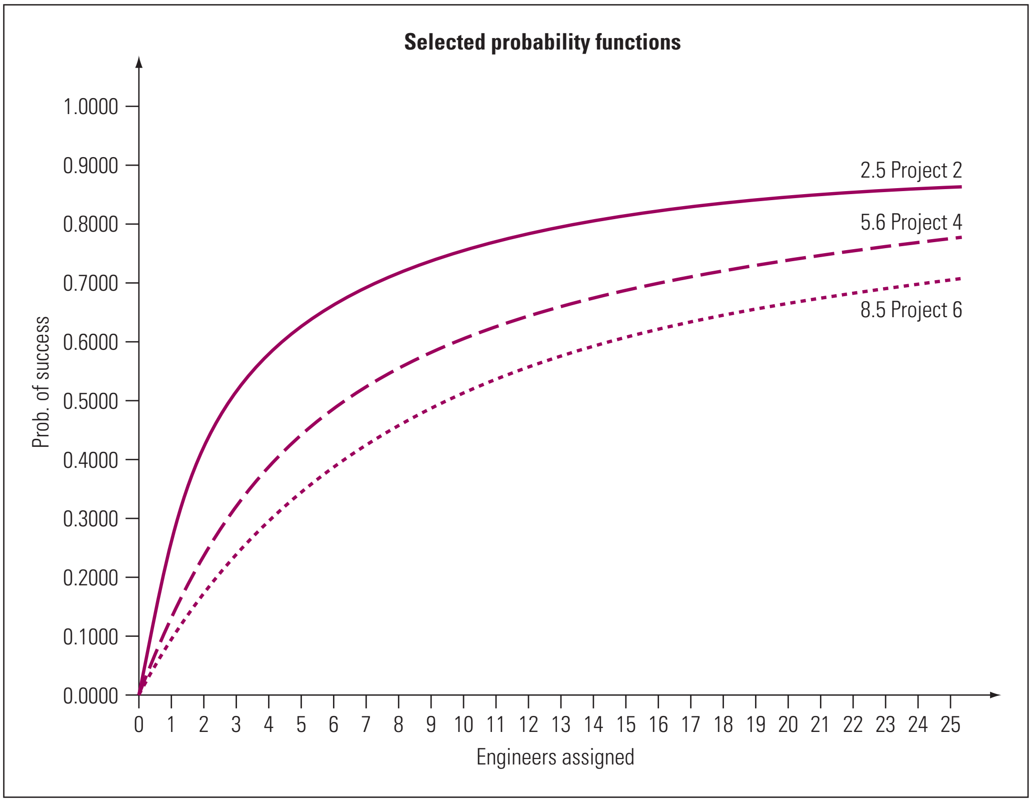


FIGURE 8.18

Graph showing the probability of success for selected projects in the TMC problem

8-7b DEFINING THE OBJECTIVE FUNCTION

TMC's directors want to maximize the expected NPV of their decision, so our objective function must correspond to this quantity. This requires that we multiply each project's expected return by the probability of the project being successful. This is accomplished as follows:

$$\text{MAX: } \frac{750X_1}{(X_1 + 3.1)} + \frac{120X_2}{(X_2 + 2.5)} + \frac{900X_3}{(X_3 + 4.5)} + \frac{400X_4}{(X_4 + 5.6)} + \frac{1100X_5}{(X_5 + 8.2)} + \frac{800X_6}{(X_6 + 8.5)}$$

8-7c DEFINING THE CONSTRAINTS

Several constraints apply to this problem. We must ensure that the projects selected require no more than \$1.7 million in startup funds. This is accomplished as follows:

$$325Y_1 + 200Y_2 + 490Y_3 + 125Y_4 + 710Y_5 + 240Y_6 \leq 1700 \quad \} \text{ Constraint on startup funds}$$

Next, we must ensure that no more than 25 engineers are assigned to selected projects. This is accomplished by the following constraint:

$$X_1 + X_2 + X_3 + X_4 + X_5 + X_6 \leq 25 \quad \} \text{ Constraint on engineers}$$

Finally, we need to make sure that engineers are assigned only to the projects that have been selected. This requires the use of linking constraints that were first presented in Chapter 6 when discussing the fixed-charge problem. The linking constraints for this problem could be stated as:

$$X_i - 25Y_i \leq 0, \quad i = 1, 2, 3, \ldots, 6 \quad \} \text{ Linking constraints}$$

These linking constraints ensure that an X_i variable can be greater than 0 if and only if its associated Y_i variable is 1.

Instead of using the previous constraint on the number of engineers and the associated six linking constraints, we could have used the following single nonlinear constraint:

$$X_1Y_1 + X_2Y_2 + X_3Y_3 + X_4Y_4 + X_5Y_5 + X_6Y_6 \leq 25 \quad \} \text{ Constraint on engineers}$$

This would sum the number of engineers assigned to selected projects. (Note that if we used this constraint we would also need to multiply each term in the objective function by its associated Y_i variable. Do you see why?) Using this single nonlinear constraint might appear to be easier than the previous seven constraints. However, when you have a choice between using linear constraints and nonlinear constraints, it is almost always better to use the linear constraints.

8-7d IMPLEMENTING THE MODEL

The model for TMC's project selection problem is summarized as:

$$\text{MAX: } \frac{750X_1}{(X_1 + 3.1)} + \frac{120X_2}{(X_2 + 2.5)} + \frac{900X_3}{(X_3 + 4.5)} + \frac{400X_4}{(X_4 + 5.6)} + \frac{1100X_5}{(X_5 + 8.2)} + \frac{800X_6}{(X_6 + 8.5)}$$

Subject to:

$325Y_1 + 200Y_2 + 490Y_3 + 125Y_4 + 710Y_5 + 240Y_6 \leq 1700$ } Constraint on startup funds

$X_1 + X_2 + X_3 + X_4 + X_5 + X_6 \leq 25$ } Constraint on engineers

$X_i - 25Y_i \leq 0, \quad i = 1, 2, 3, \ldots 6$ } Linking constraints

$X_i \geq 0$ and integer

Y_i binary

Notice that this problem has a nonlinear objective function and linear constraints. One approach to implementing this model is shown in Figure 8.19 (and in the file Fig8-19.xlsm that accompanies this book). In this spreadsheet, cells B7 through B12 are used to represent our binary Y_i variables indicating whether or not each project is selected. Cells C7 through C12 represent the X_i variables indicating the number of engineers assigned to each project.

We implemented the linking constraints for this problem by entering the LHS formulas in cells D7 through D12 as follows:

Formula for cell D7: =C7 − B7*$C$14
(Copy to cells D8 through D12.)

We will constrain these values to be less than or equal to zero. The LHS for the constraint on the number of engineers assigned to projects is implemented in cell C13 as follows:

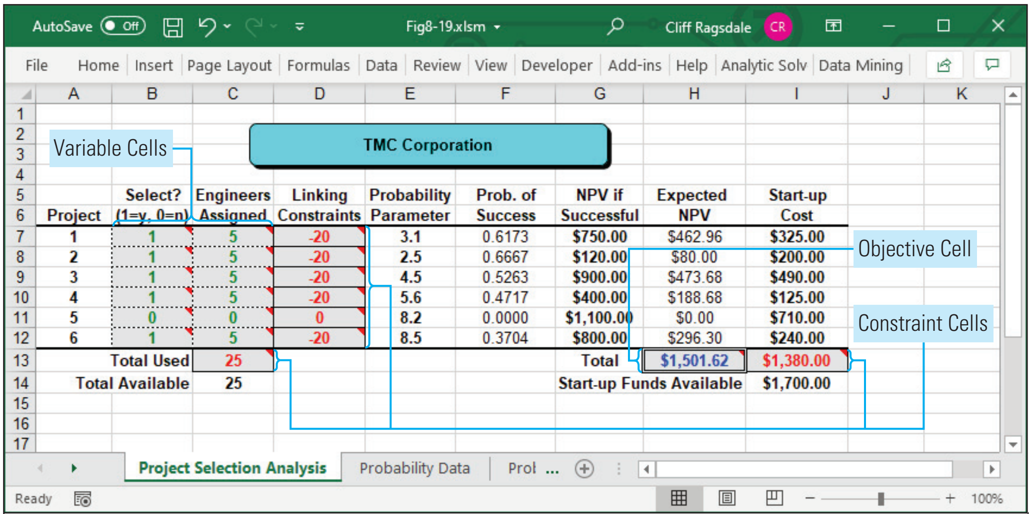
Formula for cell C13: =SUM(C7:C12)

The RHS value for this constraint is given in cell C14. Similarly, the LHS for the constraint of the total startup funds is implemented in cell I13 with its RHS value listed in I14.

Formula for cell I13: =SUMPRODUCT(B7:B12,I7:I12)

FIGURE 8.19

Spreadsheet implementation of TMC's project selection problem



Key Cell Formulas

Cell	Formula	Copied to
C13	=SUM(C7:C12)	H13
D7	=C7–B7*$C$14	D8:D12
F7	=C7/(C7+E7)	F8:F12
H7	=F7*G7	H8:H12
I13	=SUMPRODUCT(B7:B12,I7:I12)	--

To implement the objective function, we first calculate the probability of success for each project. This is done in column F as follows:

Formula for cell F7: = C7/(C7 + E7)
(Copy to cells F8 through F12.)

Next, the expected NPV value for each project is computed by multiplying the probability of success for each project by the NPV it should generate if the project is successful. This is done in column H as follows:

Formula for cell H7: = F7*G7
(Copy to cells H8 through H12.)

Finally, we compute the sum of the expected NPVs for selected projects in cell H13:

Formula for cell H13: =SUM(H7:H12)

The Solver settings and options used to solve this problem are shown in Figure 8.20.

8-7e SOLVING THE MODEL

An arbitrary starting point for this problem was selected as shown in Figure 8.19. From this arbitrary starting point, Solver located the solution shown in Figure 8.21, which has an expected

Solver Settings:

Objective: H13 (Max)
Variable cells: B7:C12
Constraints:
    C13 <= C14
    I13 <= I14
    D7:D12 <= 0
    C7:C12 >= 0
    B7:B12 = binary
    C7:C12 = integer

Solver Options:

    Standard GRG Nonlinear Engine
    Integer Tolerance = 0

FIGURE 8.20

Solver parameters for TMC's project selection problem

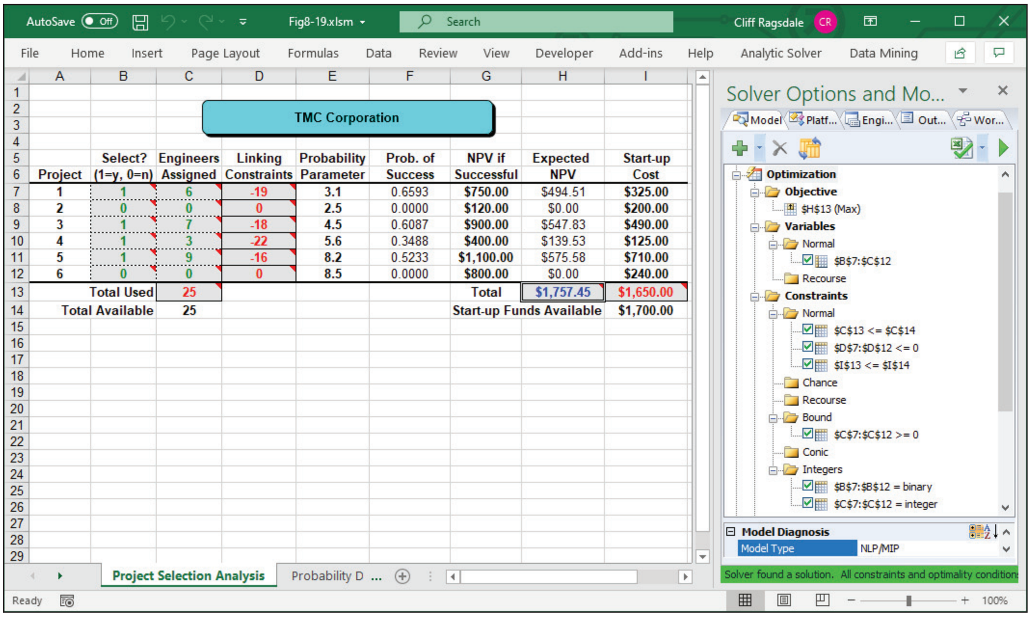


FIGURE 8.21

Solution to the TMC project selection problem

NPV of approximately $1.757 million. In this solution, notice that the probability of success for project 4 is only 0.3488. Thus, project 4 is far more likely to fail than succeed if it is assigned only three engineers. As a result, we might want to add a constraint to this problem to ensure that if a project is selected, it must have at least a 50% chance of succeeding. An exercise at the end of this chapter asks you to do that.

8-8 Optimizing Existing Financial Spreadsheet Models

So far in our discussion of optimization, we have always constructed an algebraic model of a decision problem and then implemented this model in a spreadsheet for solution and analysis. However, we can apply optimization techniques to virtually any existing spreadsheet model.

Many existing spreadsheets involve financial calculations that are inherently nonlinear. The following is an example of how optimization can be applied to an existing spreadsheet.

Thom Pearman's life is changing dramatically. He and his wife recently bought a new home and are expecting their second child in a few months. These new responsibilities have prompted Thom to think about some serious issues, including life insurance. Ten years ago, Thom purchased an insurance policy that provides a death benefit of $40,000. This policy is paid for in full and will remain in force for the rest of Thom's life. Alternatively, Thom can surrender this policy and receive an immediate payoff of approximately $6,000 from the insurance company.

Ten years ago, the $40,000 death benefit provided by the insurance policy seemed more than adequate. However, Thom now feels that he needs more coverage to care for his wife and children adequately in the event of his untimely death. Thom is investigating a different kind of insurance that would provide a death benefit of $350,000 but would also require ongoing annual payments to keep the coverage in force. He received the following estimates of the annual premiums for this new policy in each of the next 10 years:

Year	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Premium	$423	$457	$489	$516	$530	$558	$595	$618	$660	$716

To pay the premiums for this new policy, one alternative Thom is considering involves surrendering his existing policy and investing the $6,000 he would receive to generate the after-tax income needed to pay the premiums on his new policy. However, to see if this is possible, he wants to determine the minimum rate of return he would have to earn on his investment to generate after-tax investment earnings that would cover the premium payments for the new policy. Thom likes the idea of keeping the $6,000 in case of an emergency and does not want to use it to pay premiums. Thom's marginal tax rate is 28%.

8-8a IMPLEMENTING THE MODEL

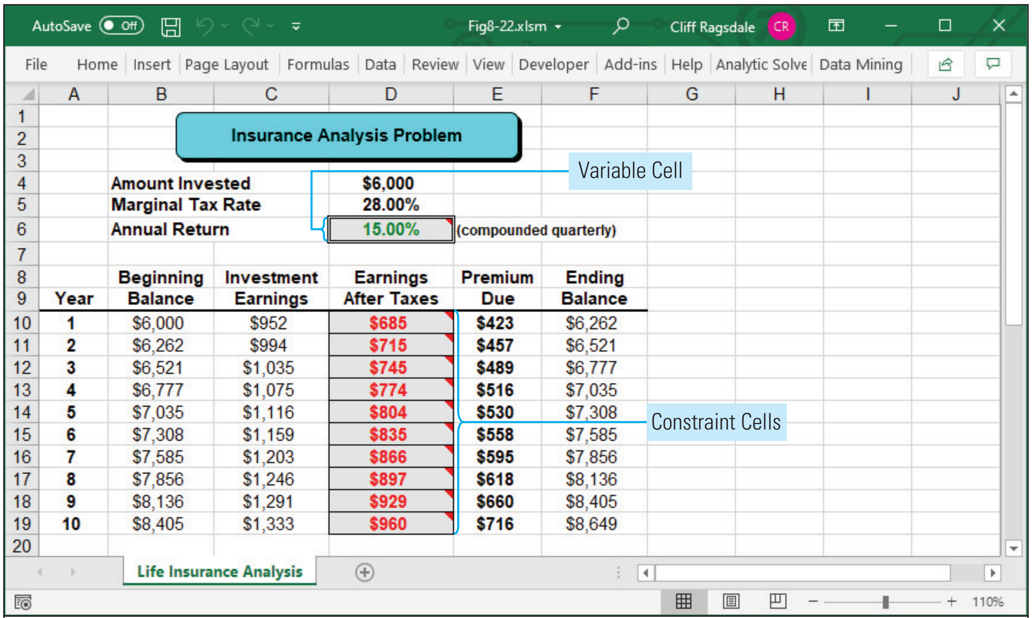
A spreadsheet model for Thom's decision problem is shown in Figure 8.22 (and in the file Fig8-22.xlsm that accompanies this book). The strategy in this spreadsheet is to determine how much money would be invested at the beginning of each year, how much money would be earned during the year after taxes, and how much would be left at the end of the year after paying taxes and the insurance premium due for that year.

As shown in Figure 8.22, cells D4 and D5 contain the assumptions about the initial amount invested and Thom's marginal tax rate. Cell D6 represents the expected annual return (which is compounded quarterly). The annual return of 15% was entered in this cell simply for planning purposes. This is the value that we will attempt to minimize when we optimize the spreadsheet.

The beginning balance for the first year (cell B10) is equal to the initial amount of money invested. The beginning balance for each of the following years is the ending balance from the previous year. The formula in cell C10 calculates the amount earned for the year given the interest rate in cell D6. This same formula applies to cells C11 through C19.

Formula for cell C10: =B10*(1 + $D$6/4)^4 − B10
(Copy to cells C11 through C19.)

Because Thom pays 28% in taxes, the values in the Earnings After Taxes column are 72% of the investment earnings listed in column C (100% − 28% = 72%). The values in the Ending Balance column are the beginning balances plus the earnings after taxes minus the premium due for the year.



Amount Invested	$6,000
Marginal Tax Rate	28.00%
Annual Return	15.00% (compounded quarterly)

Year	Beginning Balance	Investment Earnings	Earnings After Taxes	Premium Due	Ending Balance
1	$6,000	$952	$685	$423	$6,262
2	$6,262	$994	$715	$457	$6,521
3	$6,521	$1,035	$745	$489	$6,777
4	$6,777	$1,075	$774	$516	$7,035
5	$7,035	$1,116	$804	$530	$7,308
6	$7,308	$1,159	$835	$558	$7,585
7	$7,585	$1,203	$866	$595	$7,856
8	$7,856	$1,246	$897	$618	$8,136
9	$8,136	$1,291	$929	$660	$8,405
10	$8,405	$1,333	$960	$716	$8,649

Key Cell Formulas

Cell	Formula	Copied to
B10	=D4	--
B11	=F10	B12:B19
C10	=B10*(1+$D$6/4)^4-B10	C11:C19
D10	=(1-$D$5)*C10	D11:D19
F10	=B10+D10-E10	F11:F19

FIGURE 8.22

Spreadsheet implementation of Thom's insurance funding problem

8-8b OPTIMIZING THE SPREADSHEET MODEL

Three elements are involved in any optimization problem: one or more decision variables, an objective function, and constraints. The objective in the current problem is to determine the minimum annual return that will generate after-tax earnings to pay the premiums each year. Thus, the decision variable in this model is the interest rate in cell D6. The value in cell D6 also represents the objective in the problem because we want to minimize its value. For constraints, the after-tax earnings each year should be greater than or equal to the premium due for the year. Thus, we require that the values in cells D10 through D19 be greater than or equal to the values in cells E10 through E19. Figure 8.23 shows the Solver settings and options required to solve this problem, and Figure 8.24 shows the optimal solution.

8-8c ANALYZING THE SOLUTION

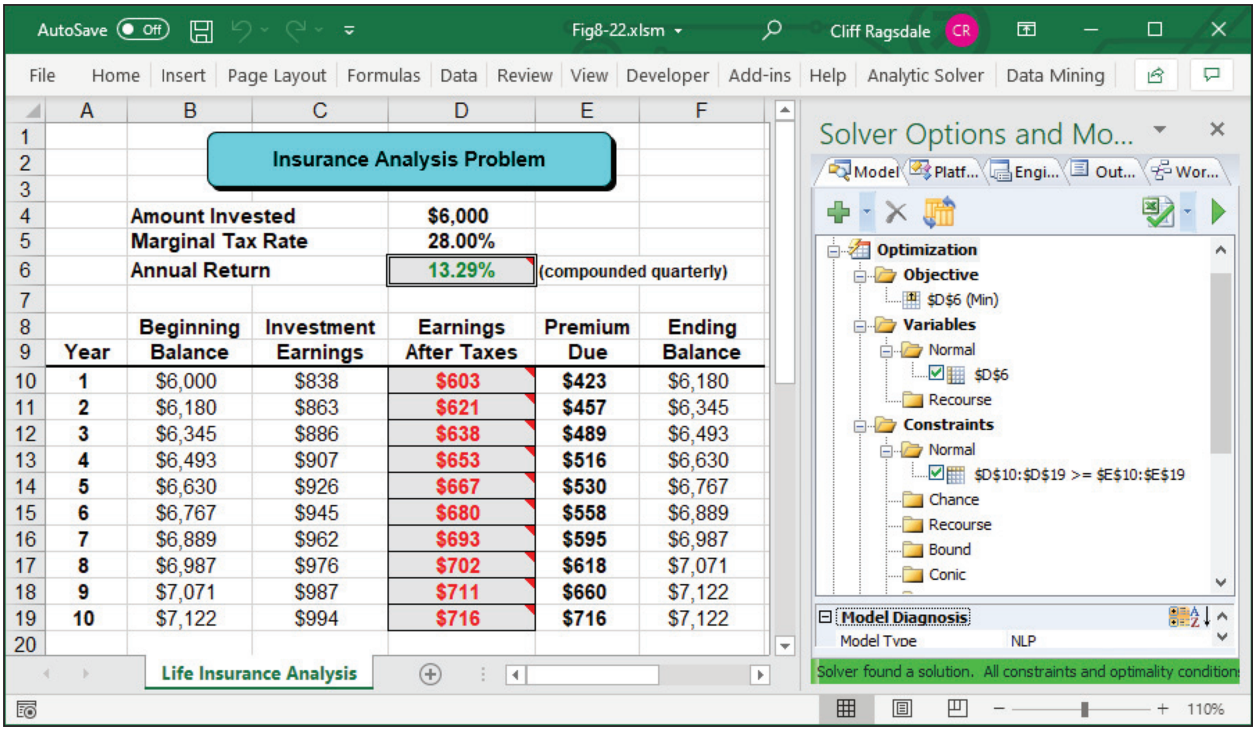
The solution shown in Figure 8.24 indicates that Thom needs to obtain an annual return of at least 13.29% in order for the after-tax earnings from his investment of $6,000 to pay the premiums on his new policy for the next 10 years. This rate of return causes his after-tax earnings in year 10 to equal exactly the premium payment of $716 due that year. Thus, in order for Thom's plan to succeed, he needs to identify an investment that is capable of producing at least a 13.29% annual return each year for the next 10 years. Thom might want to use the technique described in Section 8.9 to help design such an investment.

FIGURE 8.23

Solver parameters for the insurance funding problem

Solver Settings:
Objective: D6 (Min) Variable cells: D6 Constraints: D10:D19 >= E10:E19
Solver Options: Standard GRG Nonlinear Engine

FIGURE 8.24 *Optimal solution to the insurance funding problem*



8-8d COMMENTS ON OPTIMIZING EXISTING SPREADSHEETS

One difficulty in optimizing an existing spreadsheet model is determining whether the underlying algebraic model is linear or nonlinear. This is important in determining whether a global optimal solution to the problem has been obtained. As mentioned earlier, if we instruct Solver to assume that the model is linear, it conducts a series of numerical tests to determine whether this assumption is appropriate. If Solver detects that the model is not linear, a message is displayed to that effect and we need to re-solve the model as an NLP. So when applying optimization techniques to an existing spreadsheet model, it is a good idea to instruct Solver to assume that the model is linear. If Solver can find a solution under this assumption, we can be confident that it is the global optimal solution. If Solver detects that the model is nonlinear, we must be aware that any solution obtained might represent a local optimal solution as opposed to the global optimal solution. In this case, we might re-solve the model several times from different starting points to see if better local optimal solutions exist for the problem. (Note that if a problem is poorly scaled, Solver's linearity tests will sometimes indicate that the model is not linear when, in fact, it is.)

As you develop your skills and intuition about spreadsheet optimization, you might be inclined to skip the step of writing out algebraic formulations of your models. For straightforward problems, this might be appropriate. However, in more complex problems, this can lead to undesirable results. For example, in Chapter 6, we noted how it can be tempting to implement the binary variables in a fixed-charge problem using an IF( ) function in a spreadsheet. Unfortunately, this causes Solver to view the problem as an NLP rather than as a mixed-integer LP problem. So, how you implement the model for a problem can impact whether Solver finds the global optimal solution. As the model builder, you must understand what kind of model you have and implement it in the most appropriate way. Writing out the algebraic formulation of the model often helps to ensure that you thoroughly understand the model you are attempting to solve.

8-9 The Portfolio Selection Problem

One of the most famous applications of NLP involves determining the optimal mix of investments to hold in a portfolio in order to minimize the risk of the portfolio while achieving a desired level of return. One way to measure the risk inherent in an individual investment is the variance (or, alternatively, the standard deviation) of the returns it has generated over a period of time. One of the key objectives in portfolio selection is to smooth out the variation in the return on a portfolio by choosing investments whose returns tend to vary in opposite directions. That is, we want to choose investments that have a negative covariance or negative correlation so that when one investment generates a lower-than-average return, another investment in our portfolio offsets this with a higher-than-average return. This tends to make the variance of the return of the portfolio smaller than that of any individual investment. The following example illustrates a portfolio selection problem.

> Ray Dodson is an independent financial advisor. He recently met with a new client, Paula Ribben, who wanted Ray's advice on how best to diversify a portion of her investments. Paula has invested a good portion of her retirement savings in the stock of International Business Corporation (IBC). During the past 12 years, this stock has produced an average annual return of 7.64% with a variance of approximately 0.0026. Paula wants to earn more on her investments, but is very cautious and doesn't like to take risks. Paula has asked Ray to recommend a portfolio of investments that would provide at least a 12% average return with as little additional risk as possible. After some research, Ray identified two additional stocks, from the National Motor Corporation (NMC) and the National Broadcasting System (NBS), that he believes could help meet Paula's investment objectives. Ray's initial research is summarized in Figure 8.25 (and the file Fig8-25.xlsm that accompanies this book).
>
> As indicated in Figure 8.25, shares of NMC have produced an average rate of return of 13.43% over the past 12 years, while those of NBS have generated an average return of 14.93%. Ray used the COVAR( ) function in Excel to create the covariance matrix in this spreadsheet. The numbers along the main diagonal in this matrix correspond to the variances of the returns for each stock. For example, the covariance matrix indicates that the variances in the annual returns for IBC, NMC, and NBS are 0.00258, 0.00276, and 0.03677, respectively. The entries off the main diagonal represent covariances between different pairs of stocks. For example, the covariance between IBC and NMC is approximately −0.00025, the covariance between IBC and NBS is approximately 0.00440, and the covariance between NMC and NBS is approximately −0.00542.

FIGURE 8.25

Data for the portfolio selection problem

Portfolio Data

Year	Annual Return IBC	NMC	NBS
1	11.2%	8.0%	10.9%
2	10.8%	9.2%	22.0%
3	11.6%	6.6%	37.9%
4	-1.6%	18.5%	-11.8%
5	-4.1%	7.4%	12.9%
6	8.6%	13.0%	-7.5%
7	6.8%	22.0%	9.3%
8	11.9%	14.0%	48.7%
9	12.0%	20.5%	-1.9%
10	8.3%	14.0%	19.1%
11	6.0%	19.0%	-3.4%
12	10.2%	9.0%	43.0%
Average	7.64%	13.43%	14.93%

Covariance Matrix

	IBC	NMC	NBS
IBC	0.00258	-0.00025	0.00440
NMC	-0.00025	0.00276	-0.00542
NBS	0.00440	-0.00542	0.03677

Key Cell Formulas

Cell	Formula	Copied to
B18	=AVERAGE(B6:B17)	C18:D18
G6	=COVAR(B6:B17,$B$6:$B$17)	H6:I6
G7	=COVAR(B6:B17,$C$6:$C$17)	H7:I7
G8	=COVAR(B6:B17,$D$6:$D$17)	H8:I8

Ray wants to determine what percentage of Paula's funds should be allocated to each of the stocks in order to achieve an overall expected return of 12% while minimizing the variance of the total return on the portfolio.

8-9a DEFINING THE DECISION VARIABLES

In this problem, we must determine what percentage of the total funds invested should go toward the purchase of each of the three stocks. Thus, to formulate the model for this problem, we need the following three decision variables:

$$p_1 = \text{proportion of total funds invested in IBC}$$
$$p_2 = \text{proportion of total funds invested in NMC}$$
$$p_3 = \text{proportion of total funds invested in NBS}$$

Because these variables represent proportions, we also need to ensure that they assume values no less than 0, and that they sum to 1 (or 100%). We will handle these conditions when we identify the constraints for the problem.

8-9b DEFINING THE OBJECTIVE

The objective in this problem is to minimize the risk of the portfolio as measured by its variance. In general, the variance of a portfolio consisting of n investments is defined in most finance texts as:

$$\text{Portfolio variance} = \sum_{i=1}^{n} \sigma_i^2 p_i^2 + 2\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sigma_{ij} p_i p_j$$

where

$$p_i = \text{the percentage of the portfolio invested in investment } i$$
$$\sigma_i^2 = \text{the variance of investment } i$$
$$\sigma_{ij} = \sigma_{ji} = \text{the covariance between investments } i \text{ and } j$$

If you are familiar with matrix multiplication, you might realize that the portfolio variance can also be expressed in matrix terms as:

$$\text{Portfolio variance} = \mathbf{p}^{\mathrm{T}}\mathbf{C}\mathbf{p}$$

where

$$\mathbf{p}^{\mathrm{T}} = (p_1, p_2, \ldots, p_n)$$

$$\mathbf{C} = \text{the } n \times n \text{ covariance matrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{pmatrix}$$

Notice that if 100% of the funds available are placed in a single investment i, then the portfolio variance reduces to σ_i^2—the variance for that single investment.

In our example problem, we have:

$$\sigma_1^2 = 0.00258, \sigma_2^2 = 0.00276, \sigma_3^2 = 0.03677$$
$$\sigma_{12} = -0.00025, \sigma_{13} = 0.00440, \sigma_{23} = -0.00542$$

So, using the preceding formula, the objective for our problem is stated as:

$$\text{MIN: } 0.00258p_1^2 + 0.00276p_2^2 + 0.03677p_3^2 + 2(-0.00025p_1p_2 + 0.0044p_1p_3 - 0.00542p_2p_3)$$

This objective function is not a linear combination of the decision variables, so we must solve this problem as an NLP. However, it can be shown that the solution produced when using this objective for a portfolio selection is a global optimal solution. (This problem is actually an example of a quadratic programming [QP] problem.)

8-9c DEFINING THE CONSTRAINTS

Only two main constraints apply to this problem. As mentioned earlier, because only three investments are under consideration for this portfolio, and our decision variables represent the percentage of funds invested in each of these investments, we must ensure that our decision variables sum to 100%. This can be accomplished easily as:

$$p_1 + p_2 + p_3 = 1$$

We also need a constraint to ensure that the expected return of the entire portfolio achieves or exceeds the desired return of 12%. This condition is expressed as:

$$0.0764\, p_1 + 0.1343\, p_2 + 0.1493\, p_3 \geq 0.12$$

The LHS of this constraint represents a weighted average of the expected returns from the individual investments. This constraint indicates that the weighted average expected return on the portfolio must be at least 12%.

Finally, because the decision variables must represent proportions, we should also include the following upper and lower bounds:

$$p_1, p_2, p_3 \geq 0$$
$$p_1, p_2, p_3 \leq 1$$

The last condition, requiring each p_i to be less than or equal to 1, is mathematically redundant because the p_i must also be nonnegative and sum to 1. However, we will include this restriction for completeness.

8-9d IMPLEMENTING THE MODEL

In summary, the algebraic model for this problem is given as:

$$\text{MIN: } 0.00258p_1^2 + 0.00276p_2^2 + 0.03677p_3^2 + 2(-0.00025p_1p_2 + 0.0044p_1p_3 - 0.00542p_2p_3)$$

Subject to:

$$p_1 + p_2 + p_3 = 1$$
$$0.0764\, p_1 + 0.1343\, p_2 + 0.1493\, p_3 \geq 0.12$$
$$p_1, p_2, p_3 \geq 0$$
$$p_1, p_2, p_3 \leq 1$$

One approach to implementing this model in a spreadsheet is shown in Figure 8.26 (and in the file Fig8-26.xlsm that accompanies this book). In this spreadsheet, cells G11, H11, and I11 represent the decision variables p_1, p_2, and p_3, respectively. The initial values in these cells reflect the investor's current portfolio, which consists entirely of stock in IBC.

We can implement the objective function for this problem in a number of ways. The standard approach is to implement a formula that corresponds exactly to the algebraic form of the objective function. This is represented by:

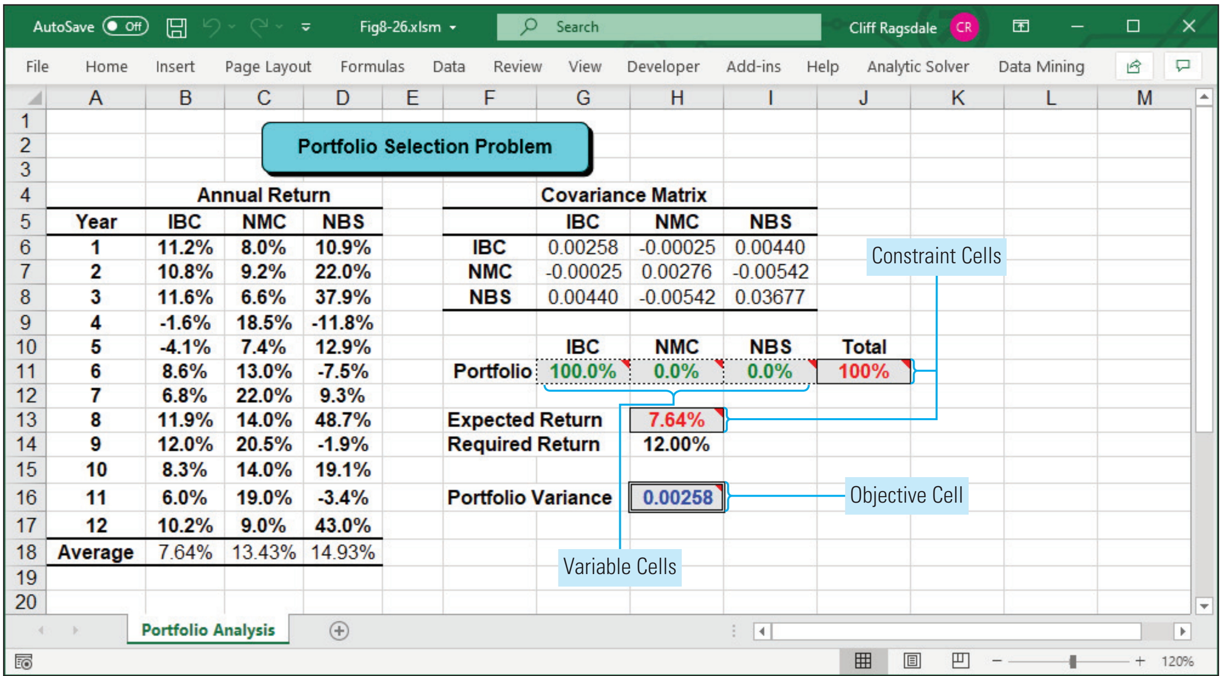
Formula for cell H16: =G6*G11^2 + H7*H11^2 + I8*I11^2+2*(H6*G11*H11 + I6*G11*I11 + H8*H11*I11)

Entering this formula is tedious and prone to error, and would be even more so if this example involved more than three stocks. The following is an alternative, and easier, way of expressing this objective:

Alternative formula for cell H16: =SUMPRODUCT(MMULT(G11:I11,G6:I8),G11:I11)

This alternative formula uses matrix multiplication (the MMULT( ) function) to compute the portfolio variance. Although both formulas generate the same result, the second formula is much easier to enter and can accommodate any number of investments. Notice that the value 0.00258 in cell H16 in Figure 8.26 indicates that when 100% of the funds are invested in IBC stock, the variance of the portfolio is the same as the variance of the IBC stock.

FIGURE 8.26 *Spreadsheet implementation for the portfolio selection problem*



Portfolio Selection Problem

	Annual Return				Covariance Matrix			
Year	IBC	NMC	NBS		IBC	NMC	NBS	
1	11.2%	8.0%	10.9%	IBC	0.00258	-0.00025	0.00440	
2	10.8%	9.2%	22.0%	NMC	-0.00025	0.00276	-0.00542	
3	11.6%	6.6%	37.9%	NBS	0.00440	-0.00542	0.03677	
4	-1.6%	18.5%	-11.8%					
5	-4.1%	7.4%	12.9%		IBC	NMC	NBS	Total
6	8.6%	13.0%	-7.5%	Portfolio	100.0%	0.0%	0.0%	100%
7	6.8%	22.0%	9.3%					
8	11.9%	14.0%	48.7%	Expected Return		7.64%		
9	12.0%	20.5%	-1.9%	Required Return		12.00%		
10	8.3%	14.0%	19.1%					
11	6.0%	19.0%	-3.4%	Portfolio Variance		0.00258		
12	10.2%	9.0%	43.0%					
Average	7.64%	13.43%	14.93%					

Key Cell Formulas

Cell	Formula	Copied to
J11	=SUM(G11:I11)	--
H13	=SUMPRODUCT(B18:D18,G11:I11)	--
H16	=SUMPRODUCT(MMULT(G11:I11,G6:I8),G11:I11)	--

The LHS formulas of the two main constraints are implemented in cells J11 and H13 as:

Formula for cell J11: =SUM(G11:I11)
Formula for cell H13: =SUMPRODUCT(B18:D18,G11:I11)

Figure 8.27 shows the Solver settings and options used to solve this problem, and Figure 8.28 shows the optimal solution.

Solver Settings:

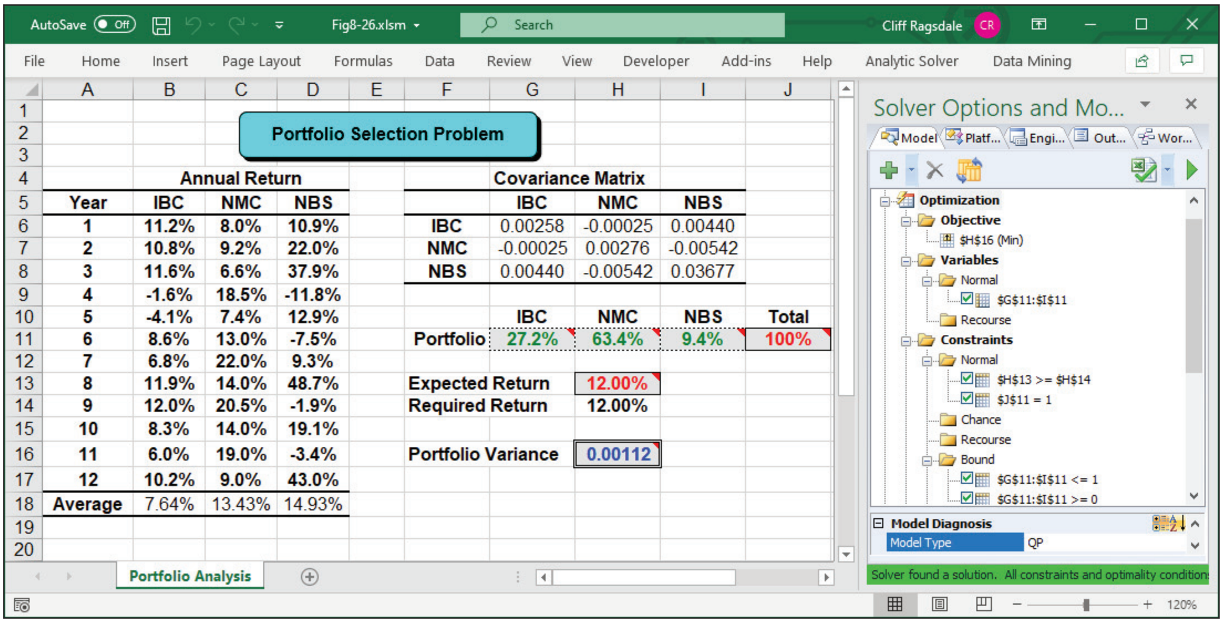
Objective: H16 (Min)
Variable cells: G11:I11
Constraints:
G11:I11 <= 1
G11:I11 >= 0
H13 >= H14
J11 = 1

Solver Options:

Standard GRG Nonlinear Engine

FIGURE 8.27 *Solver parameters for the portfolio selection problem*

FIGURE 8.28 *Optimal solution to the portfolio selection problem*



	A	B	C	D	E	F	G	H	I	J
1										
2			Portfolio Selection Problem							
3										
4			Annual Return				Covariance Matrix			
5	Year	IBC	NMC	NBS			IBC	NMC	NBS	
6	1	11.2%	8.0%	10.9%		IBC	0.00258	-0.00025	0.00440	
7	2	10.8%	9.2%	22.0%		NMC	-0.00025	0.00276	-0.00542	
8	3	11.6%	6.6%	37.9%		NBS	0.00440	-0.00542	0.03677	
9	4	-1.6%	18.5%	-11.8%						
10	5	-4.1%	7.4%	12.9%			IBC	NMC	NBS	Total
11	6	8.6%	13.0%	-7.5%		Portfolio	27.2%	63.4%	9.4%	100%
12	7	6.8%	22.0%	9.3%						
13	8	11.9%	14.0%	48.7%		Expected Return		12.00%		
14	9	12.0%	20.5%	-1.9%		Required Return		12.00%		
15	10	8.3%	14.0%	19.1%						
16	11	6.0%	19.0%	-3.4%		Portfolio Variance		0.00112		
17	12	10.2%	9.0%	43.0%						
18	Average	7.64%	13.43%	14.93%						
19										
20										

8-9e ANALYZING THE SOLUTION

In contrast to the original solution shown in Figure 8.26, the optimal solution shown in Figure 8.28 indicates that a better solution would result by placing 27.2% of the investor's money in IBC, 63.4% in NMC, and 9.4% in NBS. Cell H13 indicates that this mix of investments would achieve the desired 12% expected rate of return, and cell H16 indicates that the variance for this portfolio would be only 0.00112.

The solution to this problem indicates that a portfolio exists that produces a *higher* expected return for Paula with *less* risk than was involved in her original portfolio. Paula's original investment would be called inefficient in the terms of portfolio theory. Portfolio theory stipulates that for each possible level of investment return, there is a portfolio that minimizes risk, and accepting any greater level of risk at that level of return is inefficient. Alternatively, for each level of investment risk, there is a portfolio that maximizes the return, and accepting any lower level of return at this level of risk is also inefficient.

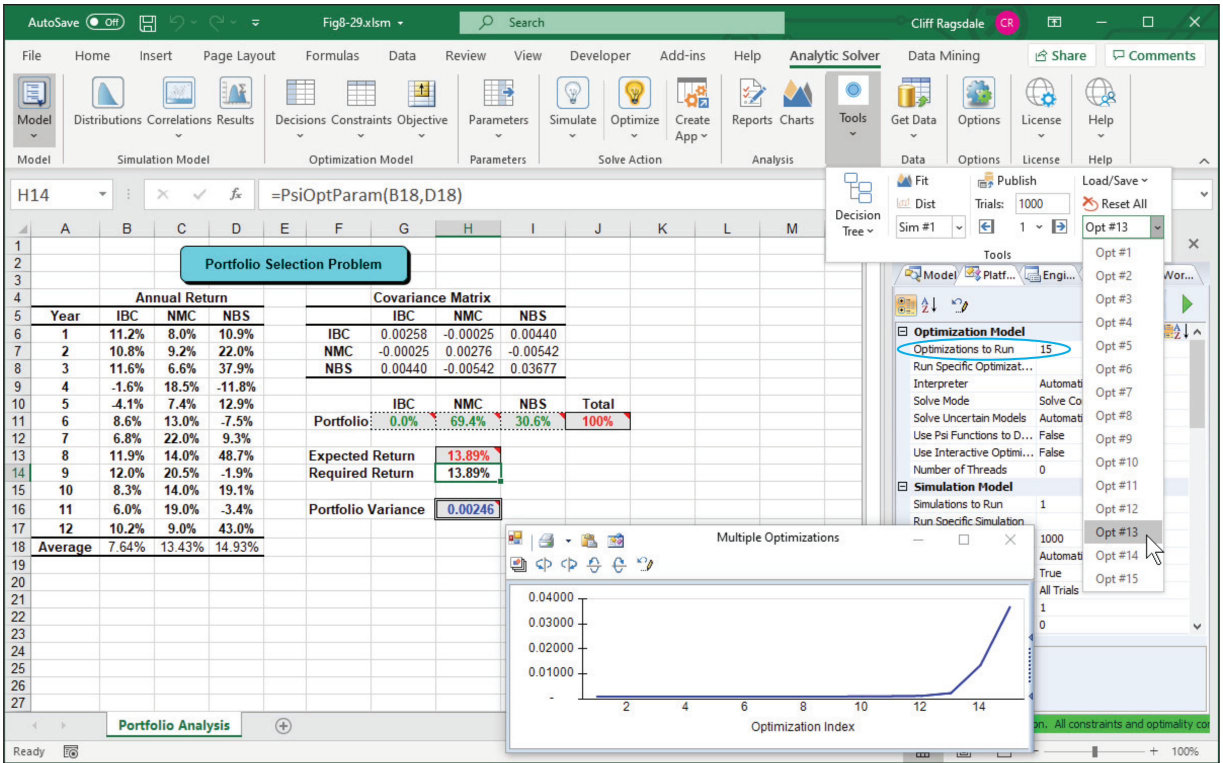
The optimal trade-off between risk and return for a given portfolio problem can be summarized by a graph of the **efficient frontier**, which plots the minimal portfolio risk associated with each possible level of return. Figure 8.29 (and the file Fig8-29.xlsm that accompanies this book) shows the efficient frontier for our example problem. This graph plots the minimal level of risk associated with 15 different portfolios where the required rate is varied in equal steps between 7.64% and 14.93% (representing, respectively, the minimum and maximum possible rates of return). To create this graph, we first placed the following formula in cell H14:

Formula for cell H14: =PsiOptParam(B18, D18)

This causes the required return value in cell H14 to be varied from 7.64% (cell B18) to 14.93% (cell D18) in equal steps as Solver runs the number of optimizations indicated by the "Optimizations to Run" setting on the Platform tab in the Solver task pane (which was set to 15 for this example). After Solver performs the optimizations, you can easily construct a graph like the one shown in Figure 8.29 by following these steps:

1. Click the Charts icon on the Analytic Solver Platform tab.
2. Select Multiple Optimizations, Monitored Cells.
3. Expand the Objective option, select $H$16, and click the ">" button.
4. Click OK.

FIGURE 8.29 *Efficient frontier for the portfolio selection problem*



Key Cell Formulas

Cell	Formula	Copied to
J11	=SUM(G11:I11)	--
H13	=SUMPRODUCT(B18:D18,G11:I11)	--
H14	=PsiOptParam(B18, D18)	--
H16	=SUMPRODUCT(MMULT(G11:I11,G6:I8),G11:I11)	--

The resulting graph in Figure 8.29 shows how the optimal portfolio variance increases for each of the 15 optimizations as the required expected return increased in equal increments from 7.64% to 14.93%. This graph is helpful not only in identifying the maximum level of risk that should be accepted for each possible level of return, but also in identifying where further increases in expected returns incur much greater amounts of risk. In this case, there is a fairly significant increase in the portfolio variance (risk) between the 13th and 14th optimization run. The drop-down list in Figure 8.29 allows us to select and inspect the details of each optimization run.

Whether you attempt to minimize risk subject to a certain required rate of return, or maximize the return subject to a given level of risk, the solutions obtained may still be inefficient. For instance, in the solution to our example problem, there may be a different portfolio that produces a higher return for the same level of risk. We could check for this by solving the problem again, maximizing the expected return while holding the minimal level of risk constant.

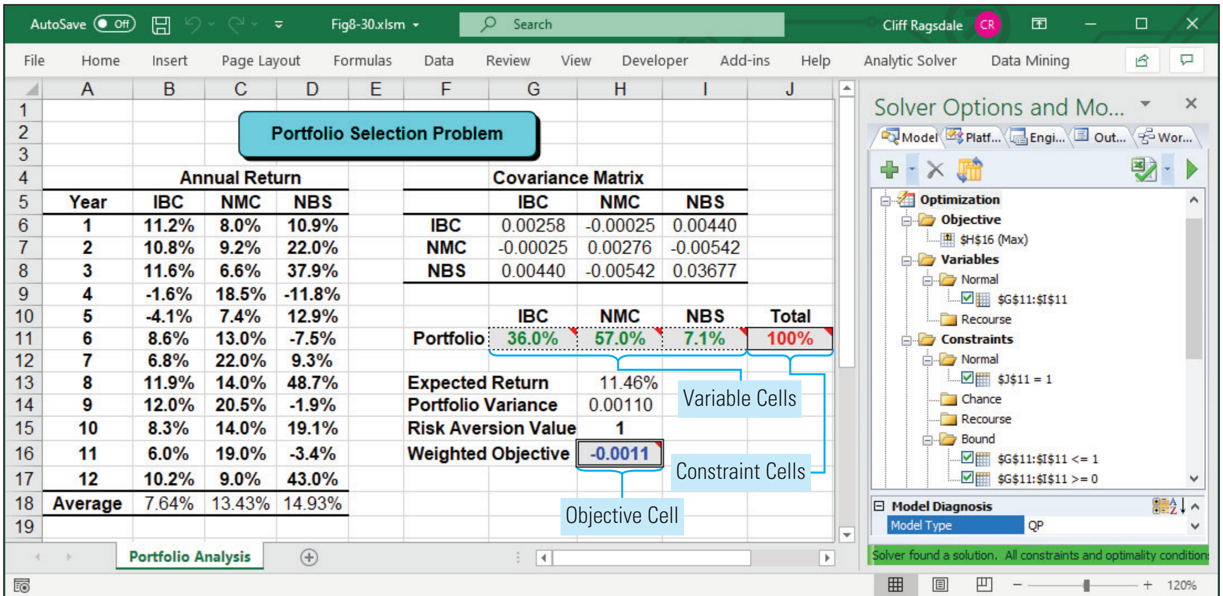
8-9f HANDLING CONFLICTING OBJECTIVES IN PORTFOLIO PROBLEMS

As we have seen, there are two different conflicting objectives that can be applied to portfolio selection problems: minimizing risk (portfolio variance) and maximizing expected returns. One way of dealing with these conflicting objectives is to solve the following problem:

$$\begin{aligned} \text{MAX:} \quad & (1 - \mathbf{r}) \times (\text{Expected Portfolio Return}) - \mathbf{r} \times (\text{Portfolio Variance}) \\ \text{Subject to:} \quad & \sum p_i = 1 \\ & p_i \geq 0 \ \text{ for all } i \end{aligned}$$

Here, the p_i again represent the percentages of money we should invest in each stock in the portfolio and **r** is a constant between 0 and 1 representing the investor's aversion to risk (or the **risk aversion value**). When $\mathbf{r} = 1$ (indicating maximum risk aversion), the objective function attempts to minimize the portfolio variance. Such a solution is shown in Figure 8.30 (and in the file Fig8-30.xlsm that accompanies this book) in which we have implemented the expected return in cell H13, the portfolio variance in cell H14, the risk aversion factor in cell H15, and

FIGURE 8.30 *Solution showing the least risky portfolio*



Year	IBC	NMC	NBS			IBC	NMC	NBS	
1	11.2%	8.0%	10.9%		IBC	0.00258	-0.00025	0.00440	
2	10.8%	9.2%	22.0%		NMC	-0.00025	0.00276	-0.00542	
3	11.6%	6.6%	37.9%		NBS	0.00440	-0.00542	0.03677	
4	-1.6%	18.5%	-11.8%						
5	-4.1%	7.4%	12.9%			IBC	NMC	NBS	Total
6	8.6%	13.0%	-7.5%		Portfolio	36.0%	57.0%	7.1%	100%
7	6.8%	22.0%	9.3%						
8	11.9%	14.0%	48.7%		Expected Return		11.46%		
9	12.0%	20.5%	-1.9%		Portfolio Variance		0.00110		
10	8.3%	14.0%	19.1%		Risk Aversion Value		1		
11	6.0%	19.0%	-3.4%		Weighted Objective		-0.0011		
12	10.2%	9.0%	43.0%						
Average	7.64%	13.43%	14.93%						

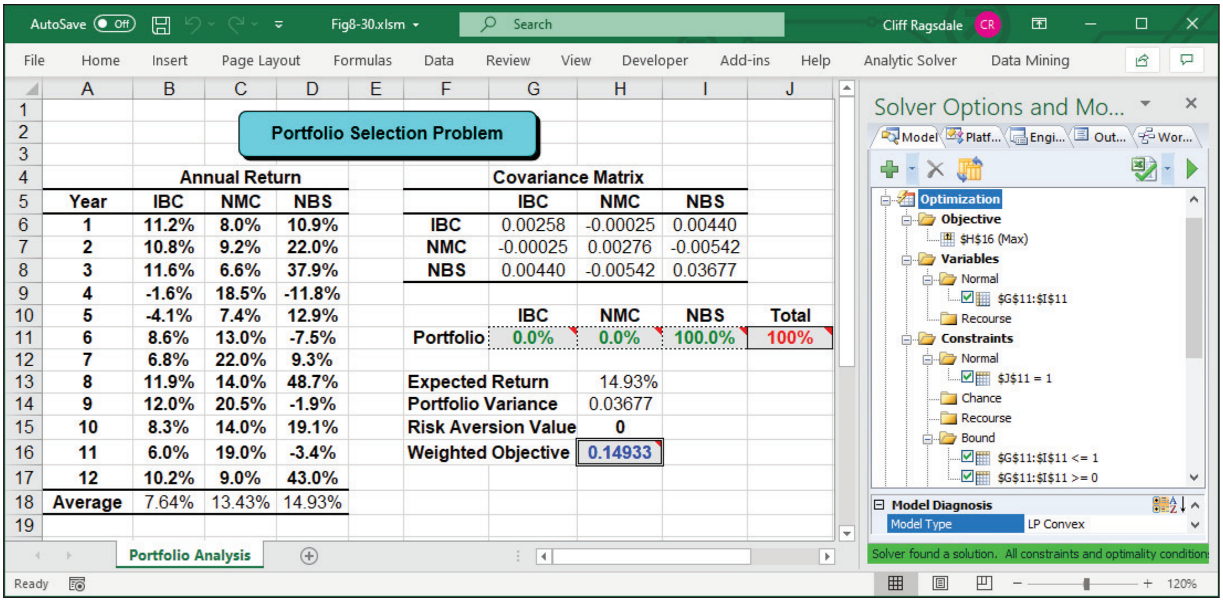
Key Cell Formulas

Cell	Formula	Copied to
J11	=SUM(G11:I11)	--
H13	=SUMPRODUCT(B18:D18,G11:I11)	--
H14	=SUMPRODUCT(MMULT(G11:I11,G6:I8),G11:I11)	--
H16	=(1-H15)*H13-H15*H14	--

the objective function in cell H16. This solution places 36% of the investor's money in IBC, 57% in NMC, and 7.1% in NBS. This results in a portfolio variance of 0.0011. This is the smallest possible portfolio variance for this collection of stocks.

Conversely, when **r** = 0 (indicating a total disregard of risk), the objective attempts to maximize the expected portfolio return. This solution is shown in Figure 8.31. This solution places 100% of the investor's money in NBS because this produces the largest return for the portfolio.

FIGURE 8.31 *Solution showing the maximum return portfolio*



For values of **r** between 0 and 1, Solver will always attempt to keep the expected return as large as possible and the portfolio variance as small as possible (because the objective function in this problem is being maximized). As the value of the parameter **r** increases, more and more weight is placed on the importance of making the portfolio variance as small as possible (or minimizing risk). Thus, a risk averse investor should prefer solutions with relatively large values of **r**. By solving a series of problems, each time adjusting the value of **r**, an investor can select a portfolio that provides the greatest utility, or the optimum balance of risk and return for their own attitudes toward risk and return. Alternatively, if an investor feels minimizing risk is twice as important as maximizing returns, we can solve the problem with **r** = 0.667 (and (1 − **r**) = 0.333) to reflect the investor's attitude toward risk and return. An **r** value of 0.99275 will produce the same solution shown earlier in Figure 8.28.

8-10 Sensitivity Analysis

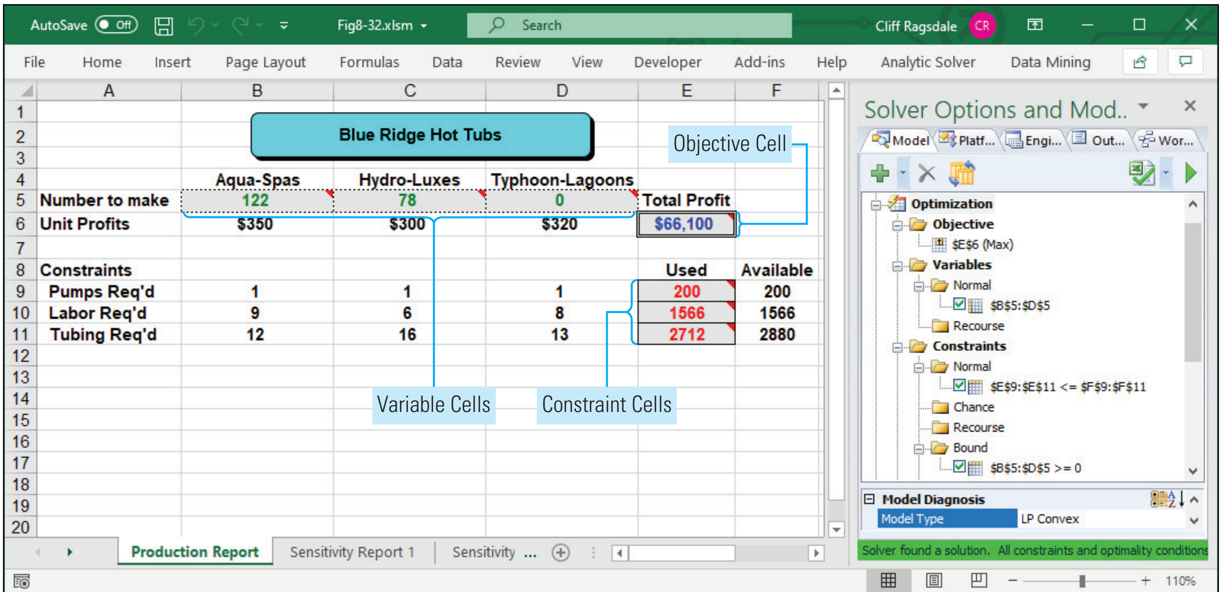
In Chapter 4, we analyzed how sensitive the optimal solution to an LP model is to changes in various coefficients in the model. We noted that one advantage of using the simplex method to solve LP problems is that it provides expanded sensitivity information. A certain amount of sensitivity information is also available when using nonlinear optimization methods to solve linear or nonlinear problems.

To understand the sensitivity information available from nonlinear optimization, we will compare it to what we learned in Chapter 4 about the sensitivity information that results from using the simplex method. In Chapter 4, we solved the following modified version of the Blue Ridge Hot Tubs problem where a third type of hot tub—the Typhoon-Lagoon—was included in the model:

$$
\begin{array}{llr}
\text{MAX:} & 350X_1 + 300X_2 + 320X_3 & \} \text{ profit} \\
\text{Subject to:} & 1X_1 + 1X_2 + 1X_3 \leq 200 & \} \text{ pump constraint} \\
 & 9X_1 + 6X_2 + 8X_3 \leq 1{,}566 & \} \text{ labor constraint} \\
 & 12X_1 + 16X_2 + 13X_3 \leq 2{,}880 & \} \text{ tubing constraint} \\
 & X_1, X_2, X_3 \geq 0 & \} \text{ nonnegativity conditions}
\end{array}
$$

The spreadsheet implementation of this problem is shown in Figure 8.32 (and in the file Fig 8-32.xlsm that accompanies this book). Figure 8.33 shows the Sensitivity Report generated for this problem after solving it using the simplex method. Figure 8.34 shows the Sensitivity Report for this problem after solving it using Solver's nonlinear optimizer.

FIGURE 8.32 *Spreadsheet model for the revised Blue Ridge Hot Tubs problem*

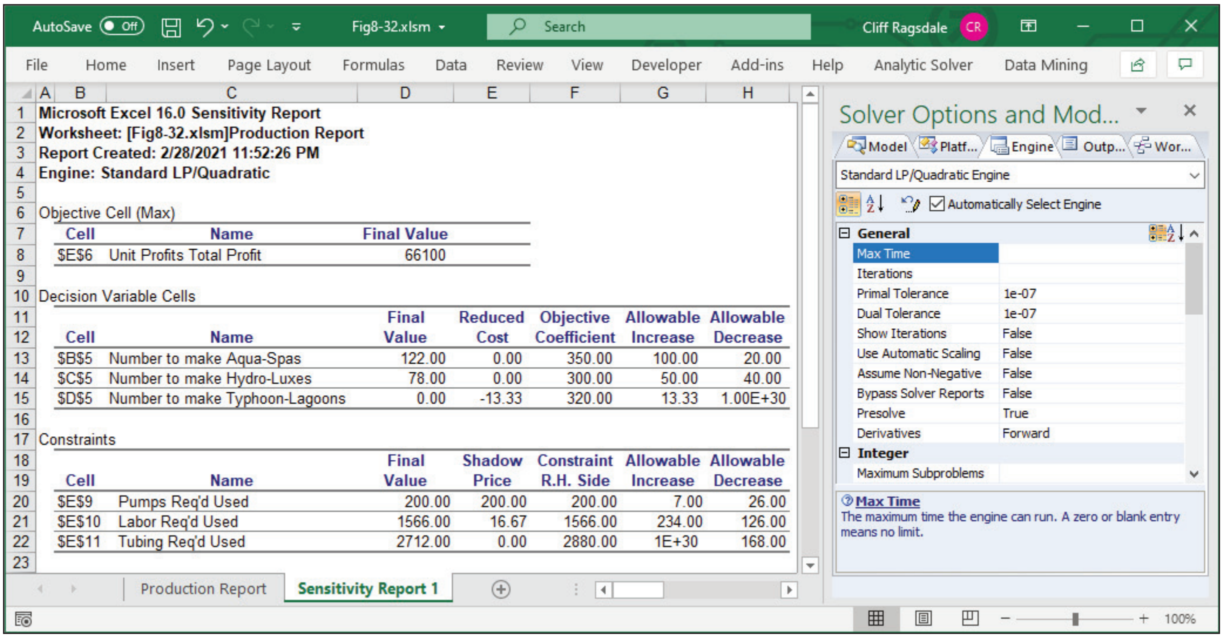


Key Cell Formulas

Cell	Formula	Copied to
E6	=SUMPRODUCT(B6:D6,$B$5:$D$5)	E9:E11

In comparing Figures 8.33 and 8.34, notice that the same optimal solution is obtained regardless of whether the problem is solved using the simplex method or the nonlinear optimizer. Both reports indicate that 122 Aqua-Spas, 78 Hydro-Luxes, and 0 Typhoon-Lagoons should be produced. Both reports also indicate that this solution requires 200 pumps, 1,566 labor hours, and 2,712 feet of tubing. The fact that the two optimization techniques found the same optimal solution is not surprising because this problem is known to have a unique optimal solution. However, if an LP problem has alternative optimal solutions, the simplex method and the nonlinear optimizer will not necessarily identify the same optimal solution.

FIGURE 8.33 *Sensitivity Report obtained after solving the model using the simplex method*



Microsoft Excel 16.0 Sensitivity Report
Worksheet: [Fig8-32.xlsm]Production Report
Report Created: 2/28/2021 11:52:26 PM
Engine: Standard LP/Quadratic

Objective Cell (Max)

Cell	Name	Final Value
$E$6	Unit Profits Total Profit	66100

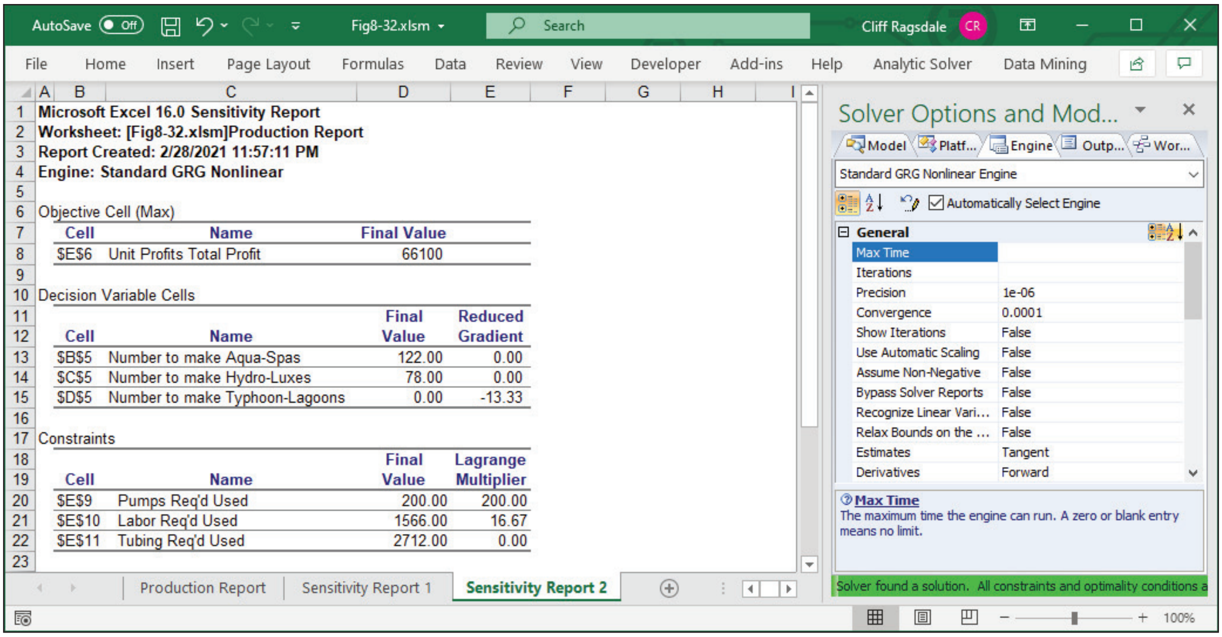
Decision Variable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient	Allowable Increase	Allowable Decrease
$B$5	Number to make Aqua-Spas	122.00	0.00	350.00	100.00	20.00
$C$5	Number to make Hydro-Luxes	78.00	0.00	300.00	50.00	40.00
$D$5	Number to make Typhoon-Lagoons	0.00	-13.33	320.00	13.33	1.00E+30

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
$E$9	Pumps Req'd Used	200.00	200.00	200.00	7.00	26.00
$E$10	Labor Req'd Used	1566.00	16.67	1566.00	234.00	126.00
$E$11	Tubing Req'd Used	2712.00	0.00	2880.00	1E+30	168.00

FIGURE 8.34 *Sensitivity Report obtained after solving the model using Solver's nonlinear GRG optimizer*



Microsoft Excel 16.0 Sensitivity Report
Worksheet: [Fig8-32.xlsm]Production Report
Report Created: 2/28/2021 11:57:11 PM
Engine: Standard GRG Nonlinear

Objective Cell (Max)

Cell	Name	Final Value
$E$6	Unit Profits Total Profit	66100

Decision Variable Cells

Cell	Name	Final Value	Reduced Gradient
$B$5	Number to make Aqua-Spas	122.00	0.00
$C$5	Number to make Hydro-Luxes	78.00	0.00
$D$5	Number to make Typhoon-Lagoons	0.00	-13.33

Constraints

Cell	Name	Final Value	Lagrange Multiplier
$E$9	Pumps Req'd Used	200.00	200.00
$E$10	Labor Req'd Used	1566.00	16.67
$E$11	Tubing Req'd Used	2712.00	0.00

Another similarity between the two Sensitivity Reports is apparent if we compare the values in the Reduced Cost and Shadow Price columns in Figure 8.33 with the values in the Reduced Gradient and Lagrange Multiplier columns in Figure 8.34. The reduced cost for each variable in Figure 8.33 is the same as the reduced gradient for each variable in Figure 8.34. Similarly, the shadow price for each constraint in Figure 8.33 is the same as the Lagrange multiplier for each constraint in Figure 8.34. This is not simply a coincidence.

8-10a LAGRANGE MULTIPLIERS

In Chapter 4, we saw that the shadow price of a constraint represents the marginal value of an additional unit of the resource represented by the constraint—or the amount by which the objective function would improve if the RHS of the constraint is loosened by one unit. This same interpretation applies in a more approximate sense to Lagrange multipliers. The main difference between shadow prices and Lagrange multipliers involves the range of RHS values over which the shadow price or Lagrange multiplier remains valid.

As discussed in Chapter 4 (and shown previously in Figure 8.33), after solving an LP problem using the simplex method, we can identify the allowable increase and decrease in a constraint's RHS value over which the shadow price of the constraint remains valid. We can do this because the objective function and constraints in an LP problem are all linear, making the impact of changes in a constraint's RHS value on the objective function value relatively easy to compute. However, in NLP problems, we have no general way to determine such ranges for the RHS values of the constraints. So, when using Solver's nonlinear optimizer to solve an optimization problem, we cannot easily determine the range of RHS values over which a constraint's Lagrange multiplier will remain valid. The Lagrange multipliers can be used only to estimate the approximate impact on the objective function of changing a constraint's RHS value by small amounts.

8-10b REDUCED GRADIENTS

In Chapter 4, we saw that the reduced cost of a variable that assumes its simple lower (or upper) bound in the optimal solution generally represents the amount by which the objective function would be reduced (or improved) if this variable were allowed to increase by one unit. Again, this same interpretation applies in a more approximate sense to reduced gradient values. In particular, nonzero reduced gradient values indicate the approximate impact on the objective function value of very small changes in the value of a given variable. For example, in Chapter 4, we saw that forcing the production of one Typhoon-Lagoon resulted in a $13.33 reduction in total profit for the problem shown in Figure 8.32. This is reflected by the reduced cost value for Typhoon-Lagoons in Figure 8.33 and the reduced gradient value for Typhoon-Lagoons in Figure 8.34.

Although we used an LP model to discuss the meaning of reduced gradients and Lagrange multipliers, their interpretation is the same for nonlinear problems. As stated earlier, an LP problem can be viewed as a special type of NLP problem where the objective function and constraints are linear.

8-11 Solver Options for Solving NLPs

Although we can represent an LP problem by a highly structured, and relatively simple, objective function and set of constraints, the objective function and constraints in an NLP problem can be virtually *any* mathematical function. Thus, it is not uncommon to encounter difficulties while trying to solve NLP problems.

Solver provides several options for controlling how it solves NLPs. These options—Estimates, Derivatives, and Search—are located in the Engine tab in the Analytic Solver task pane, as shown in Figure 8.35. The default settings for these options work well for many problems. However, if you have difficulty solving an NLP, you might try changing these options to force Solver to use a different search strategy. A complete description of these options would require an in-depth

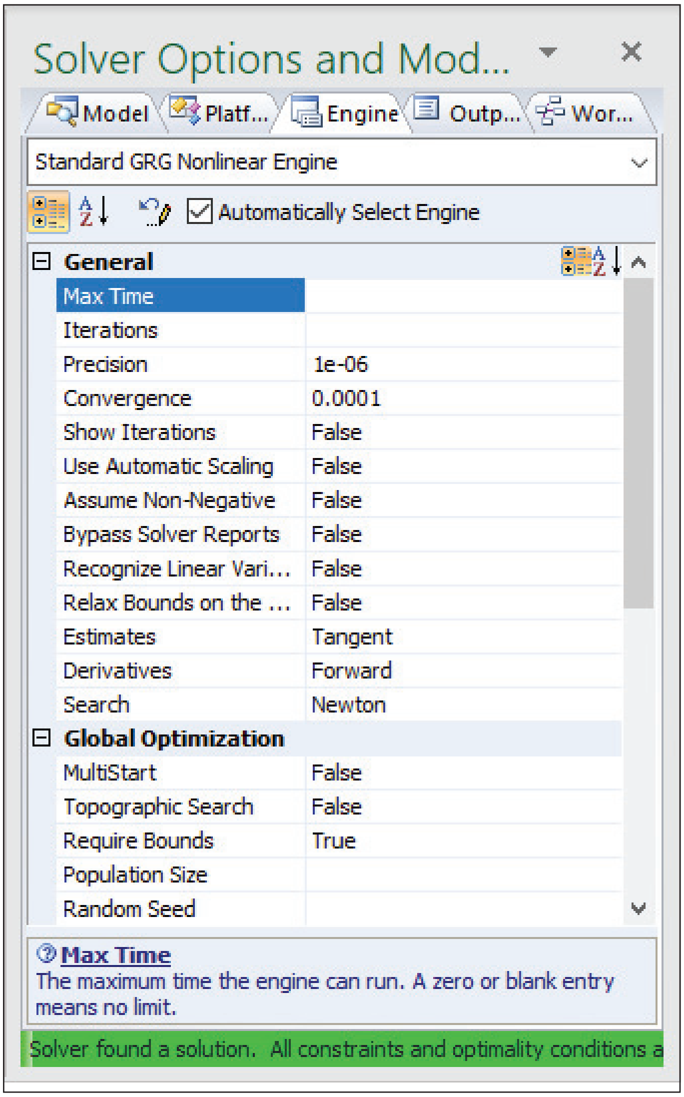


FIGURE 8.35

Solver options for NLP problems

understanding of calculus, which is not assumed in this book. The following descriptions provide a nontechnical overview of these options.

As Solver searches for an optimal solution to an NLP, it terminates if the relative change in the objective function value for several iterations is smaller than the convergence factor. If you think Solver is stopping too quickly as it converges on an optimal solution, you should reduce the convergence setting shown in Figure 8.35.

The Estimates option determines how Solver estimates the values of the decision variables while searching for improved solutions. The default setting, Tangent, estimates values using a linear extrapolation technique, whereas the alternate setting, Quadratic, uses a nonlinear extrapolation technique.

The Derivatives option determines how Solver estimates derivatives. When using the default setting, Forward, Solver obtains estimates of first derivatives at a point by perturbing the point once in a forward direction and computing the rise over the run. With the Central setting, Solver obtains estimates of first derivatives by perturbing away from a point in both a backward and forward direction and computing the rise over the run between the two points. The Central setting requires twice as many recalculations as the Forward option but can improve the estimates of the derivatives, yielding better search directions and often fewer iterations. However, the difference in accuracy is usually not worth the extra effort, hence the default is Forward.

The Search option determines how Solver chooses a search direction along which to seek a feasible point with an improved objective value. The default setting, Newton, causes Solver to use the Broyden-Fletcher-Goldfarb-Shanno Quasi-Newton method to identify search directions. The Conjugate setting instructs Solver to use the conjugate gradient method. The details of these techniques go beyond the scope of this text but can be found in most texts devoted to NLP.

As mentioned earlier, the local optimal solution at which an NLP algorithm terminates often depends on the initial starting point. Note that the MultiStart option in Figure 8.35, if set to True, causes Solver to apply methods to attempt to find a global, rather than a local, optimal solution. Additionally, scaling problems often affect how easily Solver can solve a problem. Thus, selecting the Use Automatic Scaling option is also a possible remedy to try if Solver encounters difficulty in solving an NLP.

8-12 Evolutionary Algorithms

In recent years, one of the most interesting and exciting developments in the field of optimization has centered on research into the area of evolutionary (or genetic) algorithms. Inspired by ideas from Darwin's theory of evolution, researchers interested in mathematical optimization have devised heuristic search techniques that mimic processes in biological reproduction and apply the principle of "survival of the fittest" to create general-purpose optimization engines.

In a nutshell, genetic algorithms (GAs) start with a set of chromosomes (numeric vectors) representing possible solutions to an optimization problem. The individual components (numeric values) within a chromosome are referred to as genes. New chromosomes are created by crossover and mutation. **Crossover** is the probabilistic exchange of values between solution vectors. **Mutation** is the random replacement of values in a solution vector. Chromosomes are then evaluated according to a fitness (or objective) function with the fittest surviving into the next generation. The result is a gene pool that evolves over time to produce better and better solutions to a problem.

Figure 8.36 gives an example of how one iteration through the evolutionary process might work on a simple problem involving four decision variables. In this case, we arbitrarily started with a population of 7 possible solution vectors (chromosomes). (In reality, most GAs use a population size of 50 to 100 chromosomes.) Each chromosome is evaluated according to some fitness (objective) function for the problem.

Next, we apply the crossover and mutation operators to generate new possible solutions to the problem. The second table in Figure 8.36 shows the results of this process. Note that the values for X_3 and X_4 in chromosomes 1 and 2 have been exchanged, as have the values for X_1 and X_2 in chromosomes 5 and 6. This represents the crossover operation. Also note that the values of X_2, X_3, and X_4 in chromosomes 3, 4, and 7, respectively, have been changed, randomly representing mutation. The fitness of each new chromosome is then calculated and compared against the fitness of the corresponding chromosome in the original population, with the most fit chromosome surviving into the next population. Various procedures can be used to implement the crossover, mutation, and survival of the fittest. This simple example is intended to give you a basic understanding of how a GA might work.

To a certain extent, Solver's evolutionary algorithm picks up where its nonlinear GRG algorithm leaves off. As we have seen, for nonlinear problems, the solution Solver generates depends on the starting point and may be a local rather than global optimal solution. Also, Solver tends to have difficulty solving problems with discontinuities and unsmooth landscapes, which are typical of spreadsheet models employing logical IF( ) functions and/or Lookup tables.

FIGURE 8.36

Example of one iteration through an evolutionary algorithm

Initial Population

Chromosome	X_1	X_2	X_3	X_4	Fitness
1	7.84	24.39	28.95	6.62	282.08
2	10.26	16.36	31.26	3.55	293.38
3	3.88	23.03	25.92	6.76	223.31
4	9.51	19.51	26.23	2.64	331.28
5	5.96	19.52	33.83	6.89	453.57
6	4.77	18.31	26.21	5.59	229.49
7	8.72	22.12	29.85	2.30	409.68

Crossover and Mutation

Crossover

Mutation

Chromosome	X_1	X_2	X_3	X_4	Fitness
1	7.84	24.39	**31.26**	**3.55**	334.28
2	10.26	16.36	**28.95**	**6.62**	227.04
3	3.88	**19.75**	25.92	6.76	301.44
4	9.51	19.51	**32.23**	2.64	495.52
5	**4.77**	**18.31**	33.83	6.89	332.38
6	**5.96**	**19.52**	26.21	5.59	444.21
7	8.72	22.12	29.85	**4.60**	478.93

Crossover

New Population

Chromosome	X_1	X_2	X_3	X_4	Fitness
1	7.84	24.39	31.26	3.55	334.28
2	10.26	16.36	31.26	3.55	293.38
3	3.88	19.75	25.92	6.76	301.44
4	9.51	19.51	32.23	2.64	495.52
5	5.96	19.52	33.83	6.89	453.57
6	5.96	19.52	26.21	5.59	444.21
7	8.72	22.12	29.85	4.60	478.93

Although the evolutionary algorithm cannot completely avoid the possibility of becoming trapped at a local optimal solution, its use of a randomized initial gene pool and probabilistic crossover and mutation operators make this occurrence less likely. Moreover, the evolutionary algorithm can operate on virtually any spreadsheet model—even those containing IF( ) functions, Lookup tables, and custom macro functions. We will now consider a few examples of problems where Solver's evolutionary algorithm can be applied.

8-13 Forming Fair Teams

A variety of problems exist where the goal is to form fair/balanced teams from a group of people. This can happen in amateur golf tournaments where the goal is to form teams with similar handicaps, and in civic organizations that sponsor networking events where there is a desire to ensure diversity in seating arrangements for tables. Another such problem is illustrated in the following.

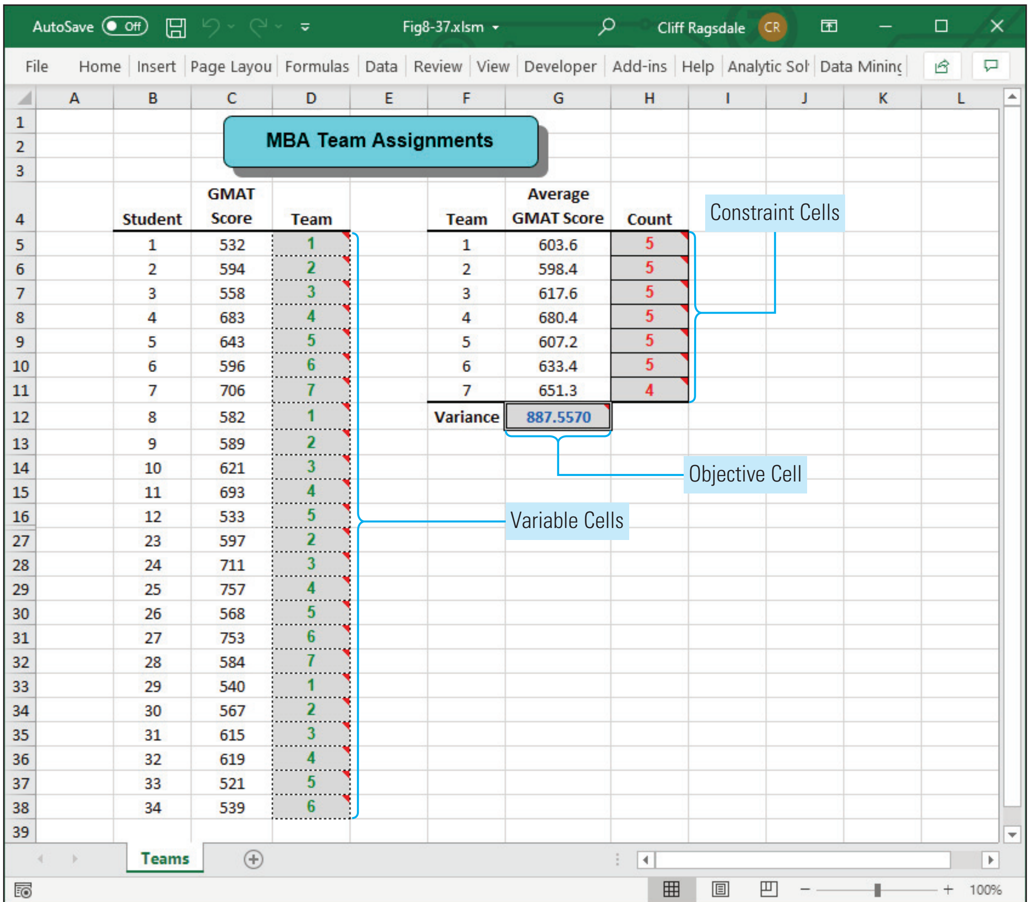
Steve Sorensen is the director of the MBA program at Claytor College. Each year, he forms project teams for the incoming class of full-time MBA students. Students work in the same team for each of their classes during their first semester in order to get to know one another and learn how to deal with people that they might not have otherwise chosen to work with. There are 34 students in the next incoming class that Steve would like to organize into seven teams. He would like to assign students to teams so that the average GMAT score for each team is as similar as possible.

8-13a A SPREADSHEET MODEL FOR THE PROBLEM

Figure 8.37 shows a spreadsheet containing GMAT scores for the 34 new MBA students for Claytor College. (Note that rows 17 to 26 of this table have been hidden to save space.) Cells D5 through D38 represent decision variables indicating to which of the seven teams each student is assigned. Arbitrary values have been assigned to these cells at present. We will instruct Solver to

FIGURE 8.37

Spreadsheet model for the MBA team assignment problem



Key Cell Formulas

Cell	Formula	Copied to
G5	=AVERAGEIF($D$5:$D$38,F5,$C$5:$C$38)	G6:G11
H5	=COUNTIF($D$5:$D$38,F5)	H6:H11
G12	=IFERROR(VAR(G5:G11),999999999)	--

assign integer values from 1 to 7 to each of these cells. Cells G5 through G11 compute the average GMAT score for students assigned to each team as follows:

Formula for cell G5: =AVERAGEIF($D$5:$D$38, F5, $C$5:$C$38)
(Copy to cells G6 through G11.)

The AVERAGEIF( ) function works in much the same way as the SUMIF( ) function covered earlier in this chapter (and in Chapter 5). In cell G5, the AVERAGEIF($D$5:$D$38, F5, $C$5:$C$38) function for team 1 compares the team assignment values in cells D5 through D38 to the value of 1 in cell F5 and, when matches occur, averages the corresponding GMAT values in cells C5 through C38.

Cell G12 computes the variance of the average GMAT scores and will serve as our objective function to be minimized for this problem.

Formula for cell G12: =IFERROR(VAR(G5:G11),999999999)

The IFERROR( ) function returns an arbitrarily large value of 999999999 if an error is ever encountered in computing the variance of G5 through G11. So if Solver happens to assign values to the decision cells that produce an error value in other computations, the objective value for such a solution will be a very large (poor) value rather than being an error value. (For instance, if Solver does not assign any students to a particular group the AVERAGEIF( ) function for that group will return a division by zero error. [Only Chuck Norris can divide by zero.])

To keep an approximately equal number of students assigned to each team we will allow a maximum of five students per team. The number of students assigned to each team is computed in cells H5 through H11 as follows:

Formula for cell H5: =COUNTIF($D$5:$D$38, F5)
(Copy to cells H6 through H11.)

8-13b SOLVING THE MODEL

In this problem, we want to use Solver to determine values for the team assignment in cells D5 through D38 that minimize the variance of the team GMAT scores in cell G12 while assigning no more than five students to each team.

Unfortunately, the AVERAGEIF( ) and COUNTIF( ) functions used in this model create discontinuities causing Solver's GRG algorithm to be fairly ineffective on this problem. Indeed, if you attempt to use Solver's GRG algorithm on this problem, it goes no further than the initial solution shown in Figure 8.37. However, if we solve the problem using Solver's evolutionary algorithm, using the settings shown in Figure 8.38, we obtain the solution shown in Figure 8.39.

Solver Settings:
Objective: G12 (Min) Variable cells: D5:D38 Constraints: H5:H11 <= 5 D5:D38 <= 7 D5:D38 >= 1 D5:D38 = integer
Solver Options: Standard Evolutionary Engine

FIGURE 8.38

Solver settings and options for the MBA team assignment problem

You might have to solve the problem several times to obtain this solution and you might even find a different (or better) solution. Because this is a non-smooth optimization problem, Solver will usually find a "good" but not necessarily global or local optimal solution.

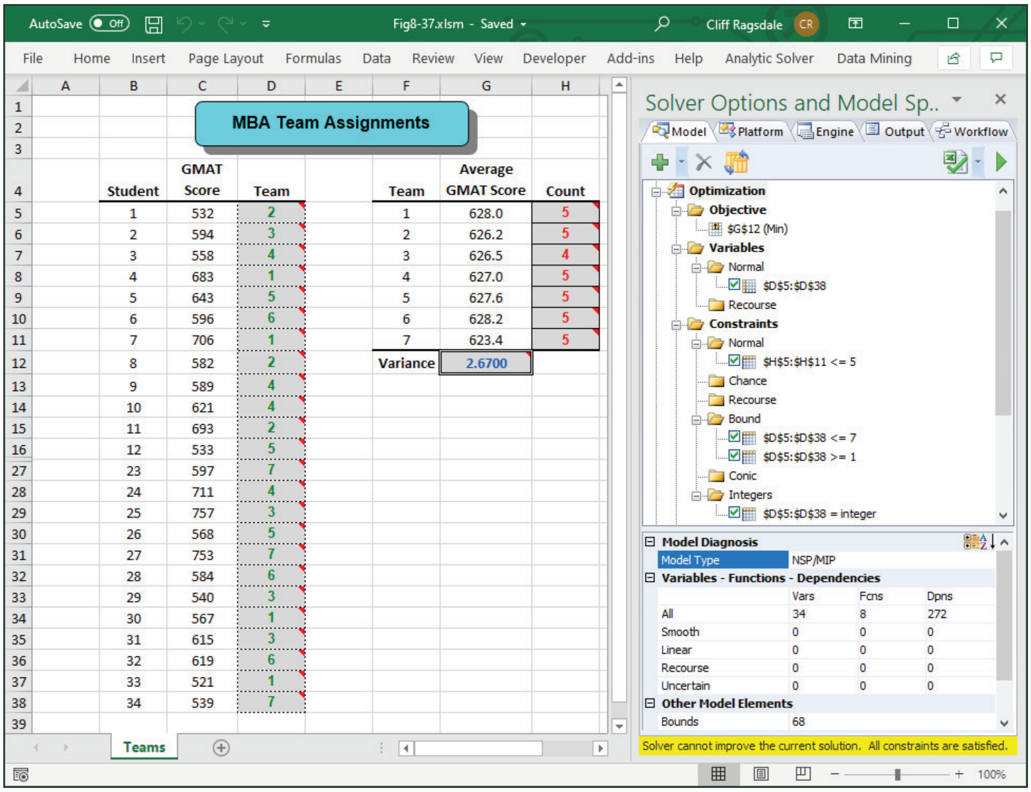


FIGURE 8.39

Possible solution for the MBA team assignment problem

8-13c ANALYZING THE SOLUTION

The team assignments shown in Figure 8.39 have reduced the variance of the average team GMAT scores significantly, from 887.5 to 2.67. It is important to remember that Solver's evolutionary algorithm is a heuristic that might or might not find the global optimal solution. Often, when Solver stops and reports that it cannot improve on the current solution, restarting Solver from a different starting point or increasing the Max Time without improvement on the Engine tab will result in a better solution.

8-14 The Traveling Salesperson Problem

The Traveling Salesperson Problem (TSP) is one of the most famous problems in the field of optimization. This problem can be described succinctly as follows:

> A salesperson wants to find the least costly (or shortest) route for visiting clients in n different cities, visiting each city exactly once before returning home.

Although this problem is very simple to state, it becomes extremely difficult to solve as the number of cities increases. In general, for an n-city TSP, there are $(n - 1)!$ possible routes (or tours) the salesman can take (where $(n-1)! = (n-1) \times (n-2) \times (n-3) \times \cdots \times 2 \times 1$). The following table shows the value of $(n-1)!$ for several different values of n:

n	$(n-1)!$
3	2
5	24
9	40,320
13	479,001,600
17	20,922,789,888,000
20	121,645,100,408,832,000

Thus, for a three-city TSP, there are only two distinct routes for the salesperson (i.e., $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$ and $1 \rightarrow 3 \rightarrow 2 \rightarrow 1$, assuming the salesperson starts in city 1). However, with just 17 cities, the number of possible routes increases to almost 21 *trillion*. Because TSPs are so difficult, heuristic solution techniques (like GAs) are often used to solve these problems.

Although it is unlikely that many traveling salespersons really care about solving this type of problem, there are numerous other examples of practical business problems that can be described in the general form of a TSP. One such example is described next.

> The Wolverine Manufacturing Company owns and operates a number of computer-controlled machines that can be programmed to perform precise drilling and machining operations. The company is currently programming their drilling machine for a job that requires nine holes to be drilled in precise locations on a flat fiberglass panel that is used in the production of a popular automobile. After each hole is drilled, the machine will automatically retract the drill bit, and move it to the next location until all the holes have been drilled. Because the machine will be required to repeat this process for millions of panels, Wolverine is interested in making sure that it programs the machine to drill the series of holes in the most efficient manner. In particular, they want to minimize the total distance the drill bit must be moved in order to complete the nine drilling operations.

If you imagine the drill bit in this problem representing a salesperson and each of the required hole locations as representing cities the drill bit must visit, it is easy to see that this is a TSP.

8-14a A SPREADSHEET MODEL FOR THE PROBLEM

To solve Wolverines' TSP, the company first must determine the straight-line distance between each pair of required hole locations. A matrix showing the distance (in inches) between each pair of required hole locations is shown in Figure 8.40 (and the file Fig8-40.xlsm that accompanies this book). Note that we are using the integers from 0 to 8 to identify the nine holes in this problem. The reason for numbering the holes starting at zero (rather than one) will become apparent shortly.

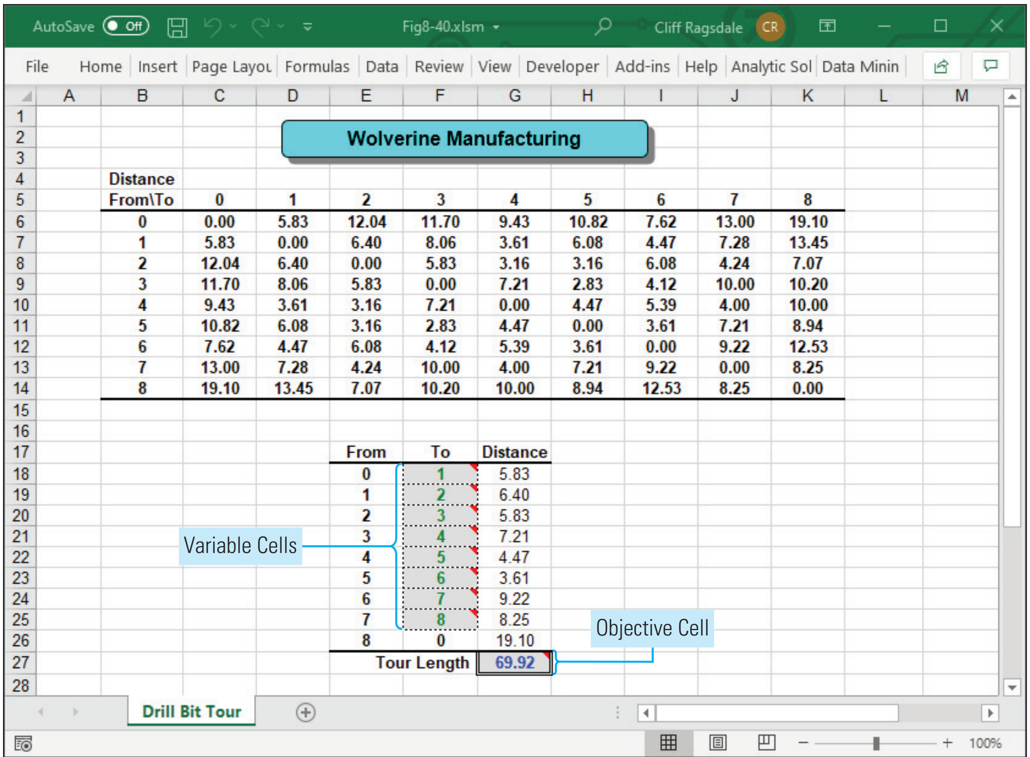
An arbitrary tour for the drill bit ($0 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6 \rightarrow 7 \rightarrow 8 \rightarrow 0$) is shown in cells E18 through F26. The distance between each of the required hole locations in this tour is shown in cells G18 through G26 using the following formula:

Formula for cell G18: =INDEX($C$6:$K$14,E18 + 1,F18 + 1)
(Copy to cells G19 through G26.)

In general, the function INDEX(***range***, ***row number***, ***column number***) returns the value in the specified ***row number*** and ***column number*** of the given ***range***. Because cell E18 contains the number 0 and F18 contains the number 1, the previous formula returns the value in the first row (E18 + 1) and second column (F18 + 1) of the range C6:K14—or the value in cell D6.

FIGURE 8.40

Spreadsheet model for Wolverine's TSP



Distance From\To	0	1	2	3	4	5	6	7	8
0	0.00	5.83	12.04	11.70	9.43	10.82	7.62	13.00	19.10
1	5.83	0.00	6.40	8.06	3.61	6.08	4.47	7.28	13.45
2	12.04	6.40	0.00	5.83	3.16	3.16	6.08	4.24	7.07
3	11.70	8.06	5.83	0.00	7.21	2.83	4.12	10.00	10.20
4	9.43	3.61	3.16	7.21	0.00	4.47	5.39	4.00	10.00
5	10.82	6.08	3.16	2.83	4.47	0.00	3.61	7.21	8.94
6	7.62	4.47	6.08	4.12	5.39	3.61	0.00	9.22	12.53
7	13.00	7.28	4.24	10.00	4.00	7.21	9.22	0.00	8.25
8	19.10	13.45	7.07	10.20	10.00	8.94	12.53	8.25	0.00

From	To	Distance
0	1	5.83
1	2	6.40
2	3	5.83
3	4	7.21
4	5	4.47
5	6	3.61
6	7	9.22
7	8	8.25
8	0	19.10
	Tour Length	69.92

Key Cell Formulas

Cell	Formula	Copied to
E19	=F18	E20:F26
G18	=INDEX($C$6:$K$14,E18+1,F18+1)	G19:G26
G27	=SUM(G18:G26)	--

Of course, any hole location that the drill bit moves *to* becomes the next location that it will move *from*. Thus, the following formula was entered in cells E19 through E26 to ensure this occurs.

Formula for cell E19: =F18
(Copy to cells E20 through E26.)

The number 0 was entered in cell F26 to ensure that the drill bit's last move is always back to its starting position. Because the solution to a TSP requires n cities to be visited exactly once, the length of the optimal tour will not change regardless of which city is selected as the starting point. (There may be alternate optimal tours, but they will all have the same objective function value.) So by selecting a starting point for a TSP, we reduce the number of possible solutions in the TSP from $n!$ to $(n-1)!$ which, as shown earlier, becomes quite significant as $\boldsymbol{n}$ increases.

The objective in this problem is to minimize the total distance the drill bit has to travel. Thus, our objective (or fitness) function should compute the total distance associated with the current tour. This is calculated in cell G27 as follows:

Formula for cell G27: = SUM(G18:G26)

If we start at hole position zero, any permutation of the set of integers from 1 to 8 in cells F18 through F25 represents a feasible tour for the drill bit. (A **permutation** is simply a rearrangement of the elements of a set.) Fortunately, Solver allows for a special type of constraint for changing cells known as the "alldifferent" constraint. The **alldifferent** constraint can be applied to a contiguous range of *n* changing cells and instructs Solver to only use a permutation of the set of integers from 1 to *n* in those cells. The "alldifferent" constraint used in combination with Solver's evolutionary optimizer allows us to model and solve a number of very challenging but practical business problems involving the optimal sequencing of jobs or activities. Several such problems are found in the questions and cases at the end of this chapter.

Solver's "alldifferent" Constraint

Solver's ***alldifferent*** constraint (selected via the "dif" option in Solver's Add Constraint dialog) can be applied to a contiguous range of *n* changing cells and instructs Solver to only use a permutation of the set of integers from 1 to *n* in those cells. Currently, you may not place any other bounds or restrictions on the cells covered by an alldifferent constraint. Thus, if you need to determine the optimal permutation of a set of integers from, for example, 21 to 28 you can:

1. Apply the alldifferent constraint to a set of eight changing cells (so Solver will generate permutations of 1 to 8 in these cells).
2. Place formulas in another set of eight cells that add 20 to the values Solver generates for the alldifferent changing cells.

8-14b SOLVING THE MODEL

Figure 8.41 shows the Solver parameters used to solve this problem. The alldifferent constraint type is selected by choosing the "dif" option when adding constraints for the variable cells. The solution obtained is shown in Figure 8.42.

FIGURE 8.41

Solver parameters for Wolverine's TSP

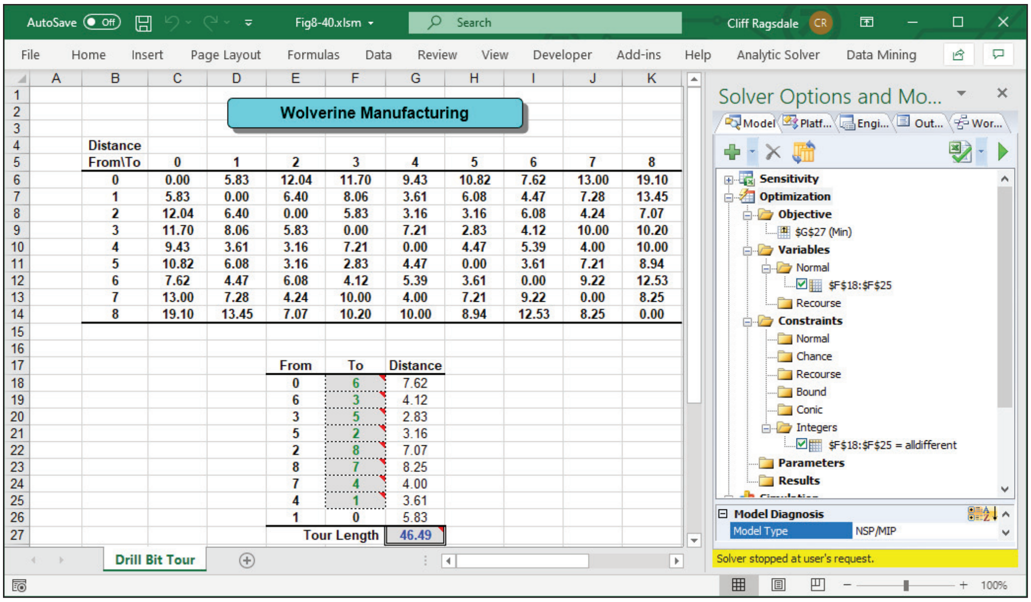
Solver Settings:
Objective: G27 (Min) Variable cells: F18:F25 Constraints: F18:F25 = alldifferent
Solver Options:
Standard Evolutionary Engine

8-14c ANALYZING THE SOLUTION

The solution shown in Figure 8.42 represents a 33.5% reduction in the total distance the drill bit needs to travel. If this drilling operation is going to be repeated on millions of parts, the reduction in processing time and machine wear and tear associated with implementing the optimal tour would likely be quite significant for this company.

FIGURE 8.42

A solution for Wolverine's TSP



It is important to remember that Solver's evolutionary algorithm randomly generates the initial population of solutions and uses probabilistic crossover and mutation operations. So if you solve this problem, you may not obtain the same solution shown in Figure 8.42 (or you may obtain an alternate optimal solution). Indeed, with large TSP type problems, if you run Solvers evolutionary optimizer several times, it will likely locate better and better solutions to the problem until the global optimal solution is found. Such is the nature of heuristic optimization techniques! As stated earlier, the evolutionary algorithm is one of the most exciting developments in the field of optimization in recent years. Solver's evolutionary search capabilities will undoubtedly continue to be refined and improved in the future.

8-15 Summary

This chapter introduced some of the basic concepts involved in NLP and discussed several applications. The steps involved in formulating and solving an NLP problem are not very different from those required to solve an LP problem—the decision variables are identified and an objective function and any constraints are stated in terms of the decision variables. Because the objective function and constraints in an NLP problem might be nonlinear, the calculations involved in solving NLP problems are different from those included in the simplex method, which is used most often to solve LP problems. NLP problems sometimes have several local optimal solutions. Thus, finding the global optimal solution to a difficult NLP problem might require re-solving the model several times using different initial starting points.

Evolutionary (or genetic) algorithms use random search techniques and the principle of survival of the fittest to solve difficult optimization problems for which linear and nonlinear optimization techniques are not suitable. Research into GAs is ongoing, but this promises to become a very useful and powerful optimization tool for business.

8-16 References

Bazaraa, M. and C. Shetty. *Nonlinear Programming: Theory and Algorithms*. New York: Wiley, 1993.

DeWitt, C., et al. "OMEGA: An Improved Gasoline Blending System for Texaco." *Interfaces*, vol. 19, no. 1, 1985.

Fylstra, D., et al. "The Design, Use, and Impact of Microsoft's Excel Solver." *Interfaces*, vol. 28, no. 5, 1998.

Holland, J. H., "Genetic Algorithms." *Scientific American*, vol. 67, no. 1, 66–72, 1992.

Kolesar, P. and E. Blum. "Square Roots Law for Fire Engine Response Distances." *Management Science*, vol. 19, 1973.

Lasdon, L. and S. Smith. "Solving Sparse Nonlinear Programs Using GRG." *ORSA Journal on Computing*, vol. 4, no. 1, 1992.

Markowitz, H. *Portfolio Selection, Efficient Diversification of Investments*. New York: Wiley, 1959.

Reeves, C. R. "Genetic Algorithms for the Operations Researcher." *INFORMS Journal on Computing*, vol. 9, no. 3, Summer 1997, pp. 231–265.

Taylor, B., L. Moore, and E. Clayton. "R&D Project Selection and Manpower Allocation with Integer Nonlinear Programming." *Management Science*, vol. 28, no. 10, 1982.

Vollman, T., W. Berry, and C. Whybark. *Manufacturing Planning and Control Systems*. Homewood, IL: Irwin, 1987.

THE WORLD OF BUSINESS ANALYTICS

Water Spilled Is Energy Lost: Pacific Gas and Electric Uses Nonlinear Optimization to Manage Power Generation

The power produced by a hydroelectric generator is a nonlinear function of the flow rate of water through the turbine and the pressure. Pressure, or head, is determined by the difference in water level upstream from the generator.

Pacific Gas and Electric Company (PG&E), the world's largest privately held utility, generates power from fossil fuels, nuclear energy, wind, solar energy, and geothermal steam, as well as hydropower. Its Sierra, Nevada Hydro System is a complex network of 15 river basins, 143 reservoirs, and 67 power plants. Stream flow peaks markedly in the spring from snow melting in the mountains, whereas demand for electric power peaks in the summer.

Water spilled from a dam cannot be used to generate power at that dam, although it can increase the head at a dam downstream and contribute to power generation there. If the water spills at a time when all of the downstream reservoirs are full, it will spill from all the dams, and its energy will be lost forever. Hydrologists at PG&E attempt to maximize the useful generation of electricity by strategically timing controlled spills to manage the levels of all reservoirs in the system and minimize wasted flow. If done effectively, this reduces the company's reliance on fossil fuel and reduces the cost of electricity to its customers.

This problem was modeled as a nonlinear program with a nonlinear objective function and linear constraints. Because many of the constraints are network flow constraints, using a network flow algorithm along with the linear terms of the objective function produced a good starting point for the NLP algorithm. A good starting point can be a critical factor in the successful use of NLP.

PG&E management confirms that the optimization system saves between $10 and $45 million annually compared to manual systems, and the California Public Utilities Commission has recommended its use to others.

Source: Ikura, Yoshiro, George Gross, and Gene Sand Hall. "PG&E's State-of-the-Art Scheduling Tool for Hydro Systems." *Interfaces*, vol. 16, no. 1, January-February 1986, pp. 65–82.

Questions and Problems

1. Can the GRG algorithm be used to solve LP problems? If so, will it always identify a corner point of the feasible region as the optimal solution (as does the simplex method)?
2. In describing the NLP solution strategy summarized in Figure 8.2, we noted that the ***fastest*** improvement in the objective function is obtained by moving from point A in a direction that is perpendicular to the level curve of the objective function. However, there are other directions that also result in improvements to the objective.
   a. How would you describe or define the set of all directions that result in improvement to the objective?
   b. How would your answer change if the level curve of the objective function at point A was nonlinear?
3. Consider an optimization problem with two variables and the constraints $X_1 \leq 5$, $X_2 \leq 5$ where both X_1 and X_2 are nonnegative.
   a. Sketch the feasible region for this problem.
   b. Sketch level curves of a nonlinear objective for this problem that would have exactly one local optimal solution that is also the global optimal solution.
   c. Redraw the feasible region and sketch level curves of a nonlinear objective for this problem that would have a local optimal solution that is not the global optimal solution.
4. Consider the following function:

$$Y = -0.865 + 8.454X - 1.696X^2 + 0.132X^3 - 0.00331X^4$$

   a. Plot this function on an X-Y graph for positive values of X from 1 to 20.
   b. How many local maximum solutions are there?
   c. How many local minimum solutions are there?
   d. Use Solver to find the maximum value of Y using a starting value of X = 2. What value of Y do you obtain?
   e. Use Solver to find the maximum value of Y using a starting value of X = 14. What value of Y do you obtain?
5. Consider the following function:

$$Y = 37.684 - 15.315X + 3.095X^2 - 0.218X^3 + 0.005X^4$$

   a. Plot this function on an X-Y graph for positive values of X from 1 to 20.
   b. How many local maximum solutions are there?
   c. How many local minimum solutions are there?
   d. Use Solver to find the minimum value of Y using a starting value of X = 3. What value of Y do you obtain?
   e. Use Solver to find the minimum value of Y using a starting value of X = 18. What value of Y do you obtain?
6. Refer to TMC's project selection problem presented in this chapter. In the solution shown in Figure 8.21, notice that the probability of success for project 4 is only 0.3488. Thus, project 4 is almost twice as likely to fail as succeed if it is assigned only three engineers. As a result, management might want to add a constraint to this problem to ensure that if a project is selected, it must have at least a 50% chance of succeeding.
   a. Reformulate TMC's problem so that if a project is selected, it must have at least a 50% chance of succeeding.
   b. Implement your model in a spreadsheet.
   c. What is the optimal solution?

7. The PENTEL Corporation manufactures three different types of computer chips. Each type of chip requires different amounts of processing time in three different departments as summarized in the following table.

	Processing Hours Required per 100 Chips			
	Chip A	Chip B	Chip C	Hours Available
Dept 1	3	2	4	10,000
Dept 2	2	4	3	9,000
Dept 3	3	4	2	11,000

The total profit for each type of chip may be described as follows:

$$\text{Chip A profit} = -0.35A^2 + 8.3A + 540$$
$$\text{Chip B profit} = -0.60B^2 + 9.45B + 1108$$
$$\text{Chip C profit} = -0.47C^2 + 11.0C + 850$$

where A, B, and C represent the number of chips produced in 100s.

a. Formulate an NLP model for this problem.
b. Implement your model in a spreadsheet and solve it.
c. What is the optimal solution?

8. A car dealership needs to determine how to allocate its $20,000 advertising budget. They have estimated the expected profit from each dollar (X) spent in four different advertising media as follows:

Medium	Expected Profit
Newspaper	$100X^{0.7}$
Radio	$125X^{0.65}$
TV	$180X^{0.6}$
Direct Mail	$250X^{0.5}$

If the company wants to spend at least $500 on each medium, how should it allocate its advertising budget in order to maximize profit?

9. The XYZ Company produces two products. The total profit achieved from these products is described by the following equation:

$$\text{Total profit} = -0.2X_1^2 - 0.4X_2^2 + 8X_1 + 12X_2 + 1500$$
$$\text{where } X_1 = \text{thousands of units of product 1}$$
$$X_2 = \text{thousands of units of product 2}$$

Every 1,000 units of X_1 require 1 hour of time in the shipping department, and every 1,000 units of X_2 require 30 minutes in the shipping department. Each unit of each product requires two pounds of a special ingredient, of which 64,000 pounds are available. Additionally, 80 hours of shipping labor are available. Demand for X_1 and X_2 is unlimited.

a. Formulate an NLP model for this problem.
b. Implement your model in a spreadsheet and solve it.
c. What is the optimal solution?

10. A traveler was recently stranded in her car in a snowy blizzard in Wyoming. Unable to drive any farther, the stranded motorist used her cell phone to dial 911 to call for help. Because the caller was unsure of her exact location, it was impossible for the emergency operator to dispatch a rescue squad. Rescue personnel brought in telecommunications experts who determined that the stranded motorist's cell phone call could be picked up by three different communications towers in the area. Based on the strength of the signal being received at each tower, they were able to estimate the distance from each tower to the caller's location.

The following table summarizes the location (X-Y coordinates) of each tower and the tower's ***estimated*** straight line (or Euclidean) distance to the caller.

Tower	X-Position	Y-Position	Estimated Distance (miles)
1	17	34	29.5
2	12	5	4.0
3	3	23	17.5

The caller's cell phone battery is quickly discharging and it is unlikely the motorist will survive much longer in the subfreezing temperatures. However, the emergency operator has a copy of Excel on her computer and believes it may be possible, with your help, to use Solver to determine the approximate location of the stranded motorist.

a. Formulate an NLP for this problem.
b. Implement your model in a spreadsheet and solve it.
c. To approximately what location should the rescue personnel be dispatched to look for the motorist?

11. Refer to the insurance problem faced by Thom Pearman discussed in section 8-8 of this chapter. Let b_i represent the balance in his investment at the beginning of year i and let r represent the annual interest rate.
    a. What is the objective function for this problem? Is it linear or nonlinear?
    b. Write out the first two constraints for this problem algebraically. Are they linear or nonlinear?
12. In the insurance problem discussed in section 8-8 of this chapter, suppose that Thom is confident that he can invest his money to earn a 15% annual rate of return compounded quarterly. Assuming a fixed 15% return, suppose he now wants to determine the minimum amount of money he must invest in order for his after-tax earnings to cover the planned premium payments.
    a. Make whatever changes are necessary to the spreadsheet and answer Thom's question.
    b. Is the model you solved linear or nonlinear? How can you tell?
13. The yield of a bond is the interest rate that makes the present value of its cash flows equal to its selling price. Assume a bond can be purchased for $975 and generates the following cash flows:

Years from Now	1	2	3	4	5
Cash Flow	$100	$120	$90	$100	$1,200

    Use Solver to determine the yield for this bond. (*Hint:* In Excel, use the NPV( ) function to compute the present value of the cash flows.) What is the yield on this bond?
14. Suppose a gift shop in Myrtle Beach has an annual demand for 15,000 units for a souvenir kitchen magnet that it buys for $0.50 per unit. Assume it costs $10 to place an order and the inventory carrying cost is 25% of the item's unit cost. Use Solver to determine the optimal order quantity if the company wants to minimize the total cost of procuring this item.
    a. What is the optimal order quantity?
    b. What is the total cost associated with this order quantity?
    c. What are the annual order and annual inventory holding costs for this solution?
15. Vijay Bashwani is organizing a charity golf tournament where teams of four players will play in a captain's choice format. The handicaps of the 40 players who have registered for the tournament are summarized in the following table. Vijay needs to create 10 teams of four players each in such a way that the total handicap of each team is as equal as possible. He would like to do this by minimizing the variance of the total handicaps of all the teams.

Player Handicaps

0	3	6	9
0	3	6	9
0	3	6	10
0	4	6	10
0	4	7	11
1	4	7	11
1	4	7	11
1	5	8	12
2	5	8	13
2	5	8	13

a. Create a spreadsheet model for this problem and solve it.
b. What are the optimal team assignments?

16. Lex Rex is an aspiring rock band composed of college friends based in Raleigh, NC. They are planning a short tour that will take them to five other college towns throughout the Mid-Atlantic region over a 10-day period. The distances between each of the cities planned for their tour are given below:

Distances Between Cities in Miles

	Home	A	B	C	D	E
Home	0	210	353	457	65	125
A	210	0	530	797	176	173
B	353	530	0	571	755	771
C	457	797	571	0	477	395
D	65	176	755	477	0	792
E	125	173	771	395	792	0

The venues where the band will perform have other acts booked on some of the dates the band will be on tour. The following table indicates (with entries of 1) the dates that venues are available in each of the cities.

Dates Available for Concerts (1 = Available, 0 = Not Available)

City	1	2	3	4	5	6	7	8	9	10
A	1	0	1	0	1	0	1	1	0	1
B	0	1	0	1	1	1	0	0	1	1
C	0	0	0	0	1	1	0	0	1	0
D	1	0	0	0	0	1	1	0	0	0
E	1	1	1	0	0	0	0	0	1	1

a. Create a spreadsheet model that can be optimized to determine the tour that minimizes the number of miles traveled. What is that tour?
b. Now use your model to determine the tour that minimizes the number of days on the road. What is that tour?
c. Now use your model to determine the tour that minimizes the amount of driving on a tour lasting 8 days. (Assume it takes 1 day to travel from any of the cities back to the band's hometown.)

17. The file InvestmentData.xlsx that accompanies this book contains data on the average returns and covariances for 15 different mutual funds. Use this data to answer the following questions:
    a. Create the efficient frontier associated with this collection of investments assuming that for each possible level of return an investor wishes to minimize risk.
    b. What portfolio has the highest expected return? What portfolio variance is associated with this portfolio?
    c. What portfolio has the smallest expected return? What portfolio variance is associated with this portfolio?
    d. Suppose an investor wanted a portfolio with an expected return of 18% using this set of investments. What portfolio would you recommend?
18. The United Delivery Service (UDS) has four vehicles that it uses to make deliveries from its depot to 36 locations in and around Moscow, Idaho. The travel times between each of the delivery locations is summarized in the file UDSData.xlsx that accompanies this book. Assume each delivery vehicle must leave from the depot and return to it at the end of its assigned delivery route. Create a spreadsheet model that can be used to determine which delivery locations should be assigned to each vehicle and the route each vehicle should travel to make its assigned deliveries.
    a. What is the optimal solution if UDS' objective is to minimize the total travel time for all vehicles?
    b. What is the optimal solution if UDS' objective is to minimize the maximum travel time for any of the vehicles?
19. SuperCity is a large retailer of electronics and appliances. The store sells three different models of TVs that are ordered from different manufacturers. The demands, costs, and storage requirements for each model are summarized in the following table:

	Model 1	Model 2	Model 3
Annual Demand	800	500	1,500
Unit Cost	\$300	\$1,100	\$600
Storage Space Required	9 sq ft	25 sq ft	16 sq ft

It costs \$60 to do the administrative work associated with preparing, processing, and receiving orders, and SuperCity assumes a 25% annual carrying cost for all items it holds in inventory. There are 3,000 square feet of total warehouse space available for storing these items, and the store never wants to have more than \$45,000 invested in inventory for these items. The manager of this store wants to determine the optimal order quantity for each model of TV.

    a. Formulate an NLP model for this problem.
    b. Implement your model in a spreadsheet and solve it.
    c. What are the optimal order quantities?
    d. How many orders of each type of TV will be placed each year?
    e. Assuming demand is constant throughout the year, how often should orders be placed?
20. The Radford hardware store expects to sell 1,500 electric garbage disposal units in the coming year. Demand for this product is fairly stable over the year. It costs \$20 to place an order for these units and the company assumes a 20% annual holding cost on inventory. The following price structure applies to Radford's purchases of this product:

	Order Quantity		
	0 to 499	500 to 999	1,000 and up
Price per Unit	\$35	\$33	\$31

So if Radford orders 135 units, it pays \$35 per unit; if it orders 650, it pays \$33 per unit; and if it orders 1,200, it pays \$31 per unit.

a. What is the most economical order quantity and total cost of this solution? (*Hint:* Solve a separate EOQ problem for each of the order quantity ranges given and select the solution that yields the lowest total cost.)
b. Suppose the discount policy changed so that Radford had to pay \$35 for the first 499 units ordered, \$33 for the next 500 units ordered, and \$31 for any additional units. What is the most economical order quantity and what is the total cost of this solution?

21. Kwane Nutimbo's family has been planning a visit to a major theme park in Florida. As the family wants to make the most of their time at the park, Kwane has collected data (in the file ParkData.xlsx accompanying this book) that includes estimated walking distances between each attraction in the park and the estimated amount of time it takes to get through each attraction. Kwane also asked each member of his family to rate the desirability of each attraction and developed a composite rating from these values. Kwane plans to spend 8 hours visiting attractions in the park (excluding the time they spend eating) and would like his family to make the best use of their time.
    a. Create a spreadsheet model to assist Kwane in maximizing his family's enjoyment of the park. How should Kwane plan to spend his time at the park?
22. Howie Jones, owner of Blue Ridge Hot Tubs, is facing a new problem. Although sale of the two hot tubs manufactured by his company (Aqua-Spas and Hydro-Luxes) have been brisk, the company is not earning the level of profits that Howie wants to achieve. Having established a reputation for high quality and reliability, Howie believes he can increase profits by increasing the prices of the hot tubs. However, he is concerned that a price increase might have a detrimental effect on demand, so Howie has engaged a marketing research firm to estimate the level of demand for Aqua-Spas and Hydro-Luxes at various prices. The marketing research firm used the technique of regression analysis (discussed in Chapter 9) to develop a model of the relationship between the prices and demand for the hot tubs. After analyzing the situation, the marketing research firm concluded that a reasonable price range for the hot tubs is between \$1,000 and \$1,500, and that within this range, Howie can expect the demand for hot tubs in the next quarter to vary with price in the following way:

$$\text{Demand for Aqua-Spas} = 300 - 0.175 \times \text{price of Aqua-Spas}$$
$$\text{Demand for Hydro-Luxes} = 325 - 0.15 \times \text{price of Hydro-Luxes}$$

    Howie determined that the costs of manufacturing Aqua-Spas and Hydro-Luxes are \$850 and \$700 per unit, respectively. Ideally, he wants to produce enough hot tubs to meet demand exactly and carry no inventory. Each Aqua-Spa requires 1 pump, 9 hours of labor, and 12 feet of tubing; each Hydro-Lux requires 1 pump, 6 hours of labor, and 16 feet of tubing. Howie's suppliers have committed to supplying him with 200 pumps and 2,800 feet of tubing. Also, 1,566 hours of labor are available for production. Howie wants to determine how much to charge for each type of hot tub and how many of each type to produce.
    a. Formulate an NLP model for this problem.
    b. Implement your model in a spreadsheet and solve it.
    c. What is the optimal solution?
    d. Which of the resource constraints are binding at the optimal solution?
    e. What values would you expect the Lagrange multipliers to take on for these constraints? (Create a Sensitivity Report for this problem to verify your answer.)
23. Carnival Confections, Inc. produces two popular southern food items, pork rinds and fried peanuts, which it sells at a local recreation area on weekends. The owners of the business have estimated their profit function on these items to be:

$$0.6p - 0.002p^2 + 0.5f - 0.0009f^2 - 0.001pf$$

Note that p is the number of packages of pork rinds produced and f is the number of packages of fried peanuts produced. Both of these items require deep frying. The company's fryer has the capacity to produce a total of 600 packages of pork rinds and/or fried peanuts. One minute of labor is required to dry and package the pork rinds, and 30 seconds are required to dry and package the peanuts. The company devotes a total of 16 hours of labor to producing these products each week.

a. Formulate an NLP model for this problem.
b. Implement your model in a spreadsheet.
c. What is the optimal solution?

24. A new mother wants to establish a college education fund for her newborn child. She wants this fund to be worth $100,000 in 18 years.
    a. If she invests $75 per month, what is the minimum rate of return she would need to earn on her investment? Assume monthly compounding. (*Hint:* Consider using the future value function FV( ) in your spreadsheet.)
    b. Suppose the mother knows of an investment that will guarantee a 12% annual return compounded monthly. What is the minimum amount she should invest each month to achieve her goal?
25. A pharmaceutical company is hiring five new salespeople to expand its sales in a western state. Pharmaceutical sales representatives do not sell directly to doctors because doctors do not purchase and distribute drugs. However, doctors do write prescriptions and it is that activity the sales representatives try to influence. The pharmaceutical company is focusing its efforts on the ten counties in the state and estimated the number of doctors in each county as follows:

County	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Doctors	113	106	84	52	155	103	87	91	128	131

Additionally, ten possible sales regions (comprising contiguous sets of counties) have been identified as follows:

	Possible Sales Region									
County	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	1		1		1					
2		1		1		1		1		
3	1		1				1			1
4		1		1					1	
5	1				1			1		
6		1				1				1
7			1				1		1	
8		1			1			1		
9				1			1		1	
10			1			1				1

For example, if a sales representative is assigned to region 1, that individual would be responsible for counties 1, 3, and 5. Each sales representative may be assigned to only a single sales region, so not all of the possible sales regions will be used. The company would like to assign their five sales representatives to these possible regions in such a way as to ensure that at least one sales representative covers each county. If regions are assigned in such a way that more than one sales representative covers the same county, the doctors within that county would be split equally among the relevant sales representatives. Additionally, the company would like to assign regions so that the total number of doctors assigned to each sales representative is as equal as possible. (Note that if exactly the same number of doctors is assigned to each sales representative, the variance of the number of doctors assigned to each sales representative would be zero.)

a. Create a spreadsheet model for this problem and use Solver's evolutionary engine to solve it.
b. What is the optimal solution?
c. What other criteria can you think of that might be relevant to the decision makers or sales representatives in this problem?

26. The Arctic Oil Company has recently drilled two new wells in a remote area of Alaska. The company is planning to install a pipeline to carry the oil from the two new wells to a transportation and refining (T&R) center. The locations of the oil wells and the T&R center are summarized in the following table. Assume a unit change in either coordinate represents 1 mile.

	X-Coordinate	Y-Coordinate
Oil well 1	50	150
Oil well 2	30	40
T&R center	230	70

Installing the pipeline is a very expensive undertaking, and the company wants to minimize the amount of pipeline required. Because the shortest distance between two points is a straight line, one of the analysts assigned to the project believes that a separate pipe should be run from each well to the T&R center. Another alternative is to run separate pipes from each well to some intermediate substation where the two lines are joined into a single pipeline that continues on to the T&R center. Arctic Oil's management wants to determine which alternative is best. Furthermore, if using the intermediate substation is best, management wants to determine where this station should be located.

a. Create a spreadsheet model to determine how many miles of pipeline Arctic Oil must install so that it runs separate pipelines from each oil well to the T&R center. How much pipe will be needed?
b. If Arctic Oil wants to build a substation, where should it be built? How much pipe is needed in this solution?
c. Which alternative is best?
d. Suppose the substation cannot be built within a 10-mile radius of the coordinates $X = 80$, $Y = 95$. (Assume the pipeline can run through this area but the substation cannot be built in the area.) What is the optimal location of the substation now and how much pipe will be needed?

27. The Rugger Corporation is a Seattle-based R&D company that recently developed a new type of fiber substrate that is waterproof and resists dirt. Several carpet manufacturers in northeast Georgia want to use Rugger as their sole supplier for this new fiber. The locations of the carpet manufacturers are summarized in the following table:

Carpet Mill Locations	X-Coordinate	Y-Coordinate
Dalton	9	43
Rome	2	28
Canton	51	36
Kennesaw	19	4

Rugger expects to make 130, 75, 90, and 80 deliveries to the carpet producers in Dalton, Rome, Canton, and Kennesaw, respectively. The company wants to build its new plant in the location that would minimize the annual shipping miles. However, Rugger also wants to be within 50 miles of each of the new customers so that it will be easy to provide on-site technical support for any production problems that may occur.

a. Formulate an NLP model for this problem.
b. Implement your model in a spreadsheet and solve it.
c. What is the optimal location for the new plant? How many annual shipping miles are associated with this solution?
d. Suppose the company wants to identify the location that minimizes the average distance to each of its customers. Where is this location and how many annual shipping miles would Rugger incur if the new plant locates there?
e. Suppose the company wants to identify the location that minimizes maximum distance to any of its customers. Where is this location and how many annual shipping miles would Rugger incur if the new plant locates there?

28. An air-ambulance service in Colorado is interested in keeping its helicopter in a central location that would minimize the flight distance to four major ski resorts. An X-Y grid was laid over a map of the area to determine the following latitude and longitude coordinates for the four resorts:

Resort	Longitude	Latitude
Bumpyride	35	57
Keyrock	46	48
Asprin	37	93
Goldenrod	22	67

a. Formulate an NLP model to determine where the ambulance service should be located in order to minimize the total distance to each resort.
b. Implement your model in a spreadsheet and solve it. Where should the ambulance service be located?
c. What other factors might affect the decision and how might you incorporate them in your model? (Consider, for example, differences in the average number of skiers and accidents at the different resorts, and the topography of the area.)

29. The Heat-Aire Company has two plants that produce identical heat pump units. However, production costs at the two differ due to the technology and labor used at each plant. The total costs of production at the plants depend on the quantity produced, and are described as:

$$\text{Total cost at plant 1: } 2X_1^2 - 1X_1 + 15$$
$$\text{Total cost at plant 2: } X_2^2 + 0.3X_2 + 10$$

Note that X_1 is the number of heat pumps produced at plant 1 and X_2 is the number of heat pumps produced at plant 2. Neither plant can make more than 600 heat pumps. Heat pumps can be shipped from either plant to satisfy demand from four different customers. The unit shipping costs and demands for each customer are summarized in the following table.

	Customer 1	Customer 2	Customer 3	Customer 4
Plant 1	\$23	\$30	\$32	\$26
Plant 2	\$33	\$27	\$25	\$24
Demand	300	250	150	400

What is the optimal production and shipping plan if management wants to meet customer demand at the lowest total cost?

a. Formulate an NLP model for this problem.
b. Implement your model in a spreadsheet and solve it.
c. What is the optimal solution?

30. Beth Dale is the Director of Development for a nonprofit organization that depends largely on charitable gifts for its operations. Beth needs to assign four different staff people to make trips to call on four possible donors. Only one staff person can call on each donor and each staff person can make only one call. Beth estimates the probability of each staff person successfully obtaining the donation from each potential giver as follows:

	Donor			
Staff	**1**	**2**	**3**	**4**
Sam	0.95	0.91	0.90	0.88
Billie	0.92	0.95	0.95	0.82
Sally	0.95	0.93	0.93	0.85
Fred	0.94	0.87	0.92	0.86

a. Formulate an NLP model to determine the assignment of staff persons to donors that maximizes the probability of receiving all donations.
b. Implement your model in a spreadsheet and solve it. What is the optimal solution?
c. Suppose it is estimated that the donations possible from donors 1, 2, 3, and 4 are for \$1 million, \$2 million, \$0.5 million, and \$0.75 million, respectively. How should Beth assign her staff if she wants to maximize the expected value of the donations received?
d. All staffers will have the least luck soliciting funds from donor number 4, so no one really wants to be assigned to this donor. Indeed, each staffer will regret not being assigned to the donor with whom they have the highest probability of success. Suppose we define the amount of this regret for each staffer by their maximum probability of success minus the probability of success for their actual assignment. What assignment of staffers to donors will minimize the maximum regret suffered by any staffer?

31. Water is delivered throughout New York City using eight main waterlines that are connected at six pumping stations as shown in Figure 8.43. The numbers on each of the arcs indicates the maximum allowable flow of water through each waterline (in 1,000s of gallons per minute).

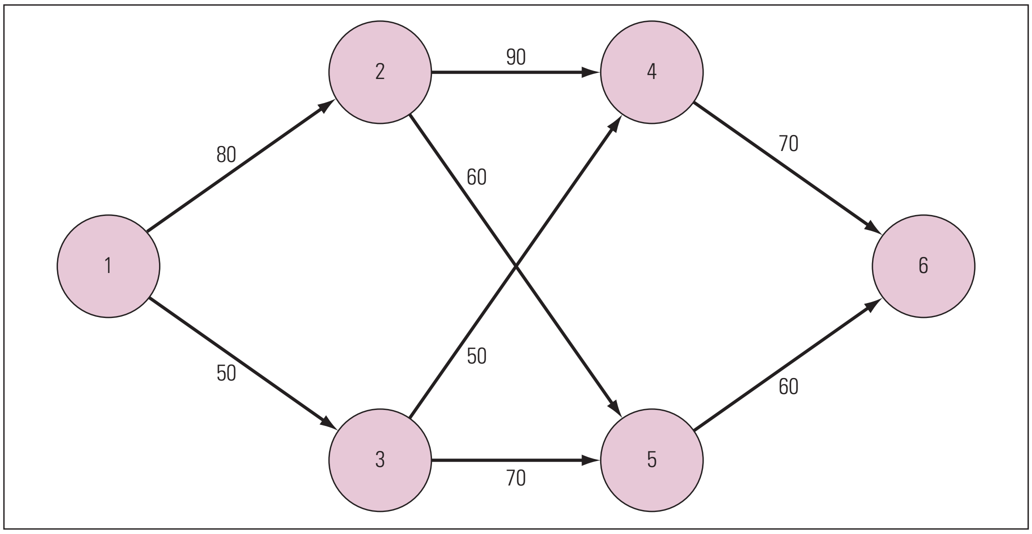


FIGURE 8.43

Main waterlines and pumping stations in New York

Because the city's waterlines are aging, breaks have been occurring more frequently and are related to the increasing demands being placed on the system. Civil engineers have estimated the probability of a waterline break occurring as follows:

Probability of failure on the line from station i to station $j = 1 - \text{EXP}(-\boldsymbol{a}_{ij}\text{F}_{ij}/1000)$

where F_{ij} is the flow (in 1,000s of gallons per minute) on the line from station i to station j and the values for the parameters $\boldsymbol{a}_{ij}$ are given as follows:

From Station	To Station	a_{ij}
1	2	0.10
1	3	0.17
2	4	0.19
2	5	0.15
3	4	0.14
3	5	0.16
4	6	0.11
5	6	0.09

Engineers can use control valves to limit how much water flows through each waterline. During peak demand times, a total of 110,000 gallons of water per minute needs to flow through this system.

a. Create a spreadsheet model to determine the flow pattern that meets the required demand for water in the most reliable way.
b. How much water should flow through each waterline?
c. What is the probability that no waterline will fail while operating in this way?

32. The Wiggly Piggly Grocery Company owns and operates numerous grocery stores throughout the state of Florida. It is developing plans to consolidate warehouse operations so that there will be 3 different warehouses that supply stores in 10 different regions of the state. The company plans to sell all its existing warehouses and build new, state-of-the-art warehouses. Each warehouse can supply multiple regions; however, all stores in a particular region will be assigned to only 1 warehouse. The locations of the different regions are summarized in the following table:

	Location	
Region	**X**	**Y**
1 Panama City	1.0	14.0
2 Tallahassee	6.1	15.0
3 Jacksonville	13.0	15.0
4 Ocala	12.0	11.0
5 Orlando	13.5	9.0
6 Tampa	11.0	7.5
7 Ft Pierce	17.0	6.0
8 Ft Myers	12.5	3.5
9 West Palm	17.5	4.0
10 Miami	17.0	1.0

a. Create a spreadsheet model to determine approximately where Wiggly Piggly should locate its new warehouses and which regions should be assigned to each of the new warehouses. Assume the company wants to build its warehouses in locations that minimize the distances to each of the regions it serves.
b. What is the optimal solution?

33. An investor wants to determine the safest way to structure a portfolio from several investments. Investment A produces an average annual return of 14% with a variance of 0.025. Investment B produces an average rate of return of 9% with a variance of 0.015. Investment C produces an average rate of return of 8% with a variance of 0.010. Investments A and B have a covariance of 0.00028, and investments A and C have a covariance of −0.006. Investments B and C have a covariance of 0.00125.
    a. Suppose the investor wants to achieve at least a 12% return. What is the least risky way of achieving this goal?
    b. Suppose the investor regards risk minimization as being five times more important than maximizing return. What portfolio would be most appropriate for the investor?
34. Muna Obiekwu wants to invest in the stocks of companies A, B, C, and D, whose annual returns for the past 13 years are as follows.

	Annual Return			
Year	**A**	**B**	**C**	**D**
1	8.0%	12.0%	10.9%	11.2%
2	9.2%	8.5%	22.0%	10.8%
3	7.7%	13.0%	19.0%	9.7%
4	6.6%	−2.6%	37.9%	11.6%
5	18.5%	7.8%	−11.8%	−1.6%
6	7.4%	3.2%	12.9%	−4.1%
7	13.0%	9.8%	−7.5%	8.6%
8	22.0%	13.5%	9.3%	6.8%
9	14.0%	6.5%	48.7%	11.9%
10	20.5%	−3.5%	−1.9%	12.0%
11	14.0%	17.5%	19.1%	8.3%
12	19.0%	14.5%	−3.4%	6.0%
13	9.0%	18.9%	43.0%	10.2%

    a. Suppose Muna is completely risk averse. What percentage of her portfolio should be invested in each stock and what would the expected risk and return be on the resulting portfolio?
    b. Suppose Muna is completely insensitive to risk and wants the maximum possible return. What percentage of her portfolio should be invested in each stock and what would the expected risk and return be on the resulting portfolio?
    c. Suppose Muna has determined her risk aversion value is **r** = 0.95. What percentage of her portfolio should be invested in each stock and what is the expected risk and return on the resulting portfolio?
35. Sometimes the historical data on returns and variances may be poor predictors of how investments will perform in the future. In this case, the ***scenario approach*** to portfolio optimization may be used. Using this technique, we identify several different scenarios describing the returns that might occur for each investment during the next year and estimate the probability associated with each scenario. A common set of investment proportions (or weights) is used to compute the portfolio return r_i for each scenario. The expected return and variance on the portfolio are then estimated as:

$$\text{Expected Portfolio Return} = EPR = \sum_i r_i s_i$$

$$\text{Variance of Portfolio Return} = VPR = \sum_i (r_i - EPR)^2 s_i$$

where r_i is the portfolio return for a given set of investment proportions under scenario i and s_i is the probability that scenario i will occur. We can use Solver to find the set of investment proportions that generate a desired *EPR* while minimizing the *VPR*. Given the following scenarios, find the investment proportions that generate an *EPR* of 12% while minimizing the *VPR*.

	Returns				
Scenario	**Windsor**	**Flagship**	**Templeman**	**T-Bills**	**Probability**
1	0.14	−0.09	0.10	0.07	0.10
2	−0.11	0.12	0.14	0.06	0.10
3	0.09	0.15	−0.11	0.08	0.10
4	0.25	0.18	0.33	0.07	0.30
5	0.18	0.16	0.15	0.06	0.40

36. Barbara Roberts recently received $30,000 as a small inheritance from a distant relative. She wants to invest the money so as to earn $900 to buy a notebook computer next year when she enters graduate school one year from now. (Barbara plans to use her inheritance to pay the tuition for her graduate studies.) She wants to create a portfolio using three stocks, whose annual percentage returns are summarized in the following table:

Year	Amalgamated Industries	Babbage Computers	Consolidated Foods
1	−3.3	5.92	−2.4
2	−4.7	−3.8	28.1
3	11.9	−7	−7.2
4	9.7	6.6	−2.3
5	8.6	−4.2	20.4
6	9.4	11.2	17.4
7	5.3	3.2	−11.8
8	−4.9	16.1	−6.6
9	8.5	10.8	−13.4
10	−8.3	−8.3	10.9

Barbara wants to diversify her potential holdings by investing at least $500, but no more than $20,000, in any one stock. Barbara wants to minimize the total variance (which also involves the covariance between the various pairs of stock) of the portfolio.

a. Formulate and solve a NLP model to determine how much money Barbara should invest in each stock in order to meet her financial goals. What is the optimal solution?
b. Will Barbara have enough money to buy a notebook computer next year?

37. World Delivery Service (WDS) specializes in the pick-up and delivery of packages at homes and businesses throughout the United States and around the world. WDS utilizes a fleet of trucks that leave from local depots and make a number of pickup and delivery stops before returning to the depot. Each such stop requires an average of three minutes of time, not including travel time between locations. Additionally, many of WDS' customers specify a specific time window within which pickups and deliveries may be made. If a WDS truck arrives at a location before the start of its time window the truck and driver must simply wait until the start of the specified time window to complete the service at that location. If the truck arrives after the close of the specified time window the customer often allows for pickup and delivery anyway, but sometimes WDS must come back the following day. Either way, arriving late creates problems for both WDS and its customers and is a practice best avoided if at all possible. The file named WDSData.xlsx that accompanies this book contains data describing travel times (in minutes) between 29 customer locations for which a given WDS

driver must make pickups and deliveries. It also lists the time windows for each customer within which these pickups and deliveries are to be made. The top priority at WDS is to provide service within its customers' specified time windows and, secondarily, to minimize the travel time of its trucks as fuel consumption is a major expense to the company. Assume the truck leaves its depot at 3:00 p.m. In what order the WDS truck service these customers?

a. Create a spreadsheet model for this problem and solve it using Solver's evolutionary engine.
b. In what order should the WDS truck service these customers?
c. How many time windows are violated in the solution you identify?
d. What is the total travel time associated with this solution?

38. A mortgage company owns the following 10 mortgages. Investors will purchase packages of mortgages worth at least $1 million. What is the maximum number of such packages that can be created from this set of mortgages?

	Mortgage									
	1	2	3	4	5	6	7	8	9	10
Amount (in $1000s)	$900	$860	$780	$525	$240	$185	$165	$164	$135	$125

a. Create a spreadsheet model for this problem and use Solver's evolutionary algorithm to solve it.
b. What is the optimal solution?

39. A small printing shop has 10 jobs it must schedule. The processing times and due dates for each job are summarized in the following table:

Job	Processing Time (Days)	Due Date
1	10	12
2	11	35
3	7	20
4	5	27
5	3	23
6	7	36
7	5	40
8	5	40
9	12	55
10	11	47

a. Suppose the jobs are scheduled in ascending order by processing time. How many jobs will be late? By how many total days will the jobs be late? What is the maximum amount by which any job is late?
b. Suppose the jobs are scheduled in ascending order by due date. How many jobs will be late? By how many total days will the jobs be late? What is the maximum amount by which any job is late?
c. Use Solver's evolutionary algorithm to determine the schedule that minimizes the number of jobs that are late. What is the solution? (Note that you may want to run Solver several times.)
d. Use Solver's evolutionary algorithm to determine the schedule that minimizes the total number of days by which the jobs are late. What is the solution?
e. Use Solver's evolutionary algorithm to determine the schedule that minimizes the maximum number of days by which any job is late. What is the solution?

40. The Major Motors Corporation manufactures heavy trucks at a plant in Dublin, Virginia. The factory's stock of spare and custom parts is stored in a huge shelving system that is several stories high and runs the length of several football fields. An automated "cherry picking" vehicle runs back and forth along a shelving unit and is able to simultaneously raise or lower to any height to pick needed stock items from the various bins in the shelving unit. Each bin in the shelving unit is of equal size and is identified by a specific row and column number. Typically, each run the cherry picker makes involves visiting several different bins to retrieve various parts. To help minimize operating costs, the company wants to develop a system to determine the most efficient way for the cherry picker to visit each required bin site before returning to its initial position. As an example, suppose the cherry picker needs to retrieve 10 parts stored in the following bin locations:

Part	Row	Column
1	3	27
2	14	22
3	1	13
4	20	3
5	20	16
6	28	12
7	30	31
8	11	19
9	7	3
10	10	25

Assume the cherry picker must start and finish at row 0 and column 0.

a. Use a spreadsheet to compute the straight-line distance between each pair of bin locations.
b. Use Solver's evolutionary algorithm to determine the shortest tour for the cherry picker to follow.
c. What is the best tour you can find?

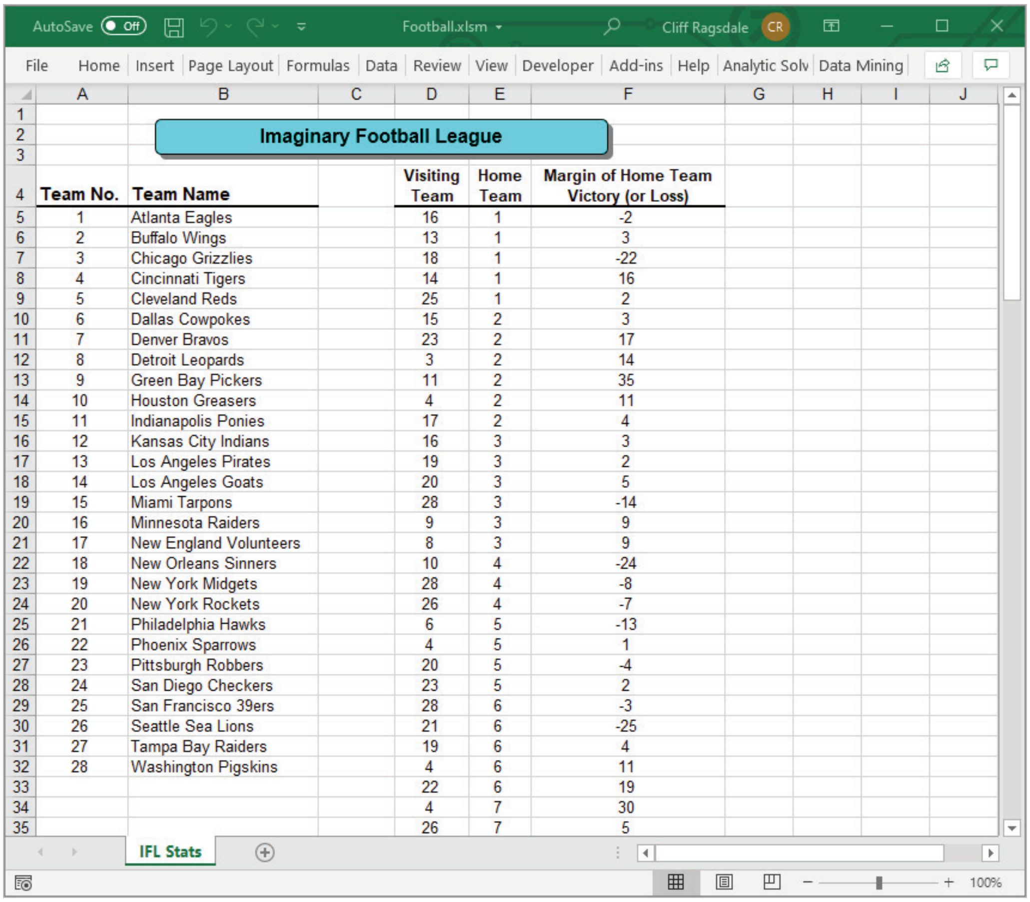
41. A regional quality inspector for Green Roof Inns has 16 properties she must visit next month. The driving time from one property to the next is proportional to the straight-line distance between the properties. The X and Y coordinate for each property are given in the following table.

	Property															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
X	190	179	170	463	153	968	648	702	811	305	512	481	763	858	517	439
Y	158	797	290	394	853	12	64	592	550	538	66	131	289	529	910	460

Assume the inspector's home is located at X coordinate 509 and Y coordinate 414.

a. Create a distance matrix that computes the straight-line distance between each pair of properties (including the inspector's home). Round these distances to two decimal places.
b. Suppose the inspector starts at her home and visits each property in numerical order before returning home. How far would she have traveled?
c. Suppose the inspector wants to start from her home and visit each property before returning home and wants to do so traveling the least distance possible. Which route should she take and how far will she travel?
d. Suppose the inspector wants to visit all 16 properties over a four week period visiting exactly 4 properties each week. Each week, she will leave from her home on Monday morning and return to her home on Friday evening. Which properties should she visit each week and in what order should she visit them if she wants to minimize the total distance she must travel?

42. Companies are often interested in segmenting their customers to better target specific product offerings to meet specific customer needs. The file CustomerData.xlsx that accompanies this book contains data on 198 customers for an online retailer. Specifically, this file lists demographic data for each customer's income level (X_1) and number of dependents (X_2) as well as buying behavior data, including the number of purchases made last year (X_3) and the average value of each purchase (X_4). Suppose you have been asked to cluster each of these customers to one of three groups. After your group assignments are made, you can compute the average values on each of the four variables within each group. These four average values for each group would represent the typical (or average) customer found in each respective group. Obviously, you want to group similar customers together. To do so, you could generalize the straight-line distance measure to four dimensions to calculate each customer's distance to his or her assigned group.
    a. Use Solver to make group assignments that minimize the sum of the distance from each customer to his or her assigned group.
    b. How would you describe the differences in the three groups or clusters you identify?
43. Max Gooding is tired of losing money in his office's weekly football pool and has decided to try to do something about it. Figure 8.44 (and file Football.xlsx that accompanies this book) contains a listing of the teams in the Imaginary Football League (IFL) along with the outcomes of all the games played in the league last season.



Team No.	Team Name	Visiting Team	Home Team	Margin of Home Team Victory (or Loss)
1	Atlanta Eagles	16	1	-2
2	Buffalo Wings	13	1	3
3	Chicago Grizzlies	18	1	-22
4	Cincinnati Tigers	14	1	16
5	Cleveland Reds	25	1	2
6	Dallas Cowpokes	15	2	3
7	Denver Bravos	23	2	17
8	Detroit Leopards	3	2	14
9	Green Bay Pickers	11	2	35
10	Houston Greasers	4	2	11
11	Indianapolis Ponies	17	2	4
12	Kansas City Indians	16	3	3
13	Los Angeles Pirates	19	3	2
14	Los Angeles Goats	20	3	5
15	Miami Tarpons	28	3	-14
16	Minnesota Raiders	9	3	9
17	New England Volunteers	8	3	9
18	New Orleans Sinners	10	4	-24
19	New York Midgets	28	4	-8
20	New York Rockets	26	4	-7
21	Philadelphia Hawks	6	5	-13
22	Phoenix Sparrows	4	5	1
23	Pittsburgh Robbers	20	5	-4
24	San Diego Checkers	23	5	2
25	San Francisco 39ers	28	6	-3
26	Seattle Sea Lions	21	6	-25
27	Tampa Bay Raiders	19	6	4
28	Washington Pigskins	4	6	11
		22	6	19
		4	7	30
		26	7	5

FIGURE 8.44

Spreadsheet for the Imaginary Football League (IFL)

For instance, cell F5 indicates the Minnesota Raiders beat the Atlanta Eagles by two points last year, whereas cell F6 indicates that Atlanta beat the Los Angeles Pirates by three points. Max believes it may be possible to use the evolutionary algorithm in Solver to estimate the margin of victory in a match-up between any two teams. Using last season's data, Max wants to identify ratings or weights for each team such that the estimated margin of victory for any match-up would be:

$$\begin{matrix}\text{Estimted Margin} \\ \text{of Victory}\end{matrix} = \left(\begin{matrix}\text{Home Team} \\ \text{Rating}\end{matrix}\right) + \left(\begin{matrix}\text{Home Field} \\ \text{Advantage}\end{matrix}\right) - \left(\begin{matrix}\text{Visiting Team} \\ \text{Rating}\end{matrix}\right)$$

With this information, Max could estimate the margin of victory in a match-up between any two teams. He wants to do this in a way that minimizes the sum of the squared differences between the actual and estimated margins of victory for each of the games last year. (Assume each team's rating should be between 0 and 100 and the home field advantage should be between 0 and 20.)

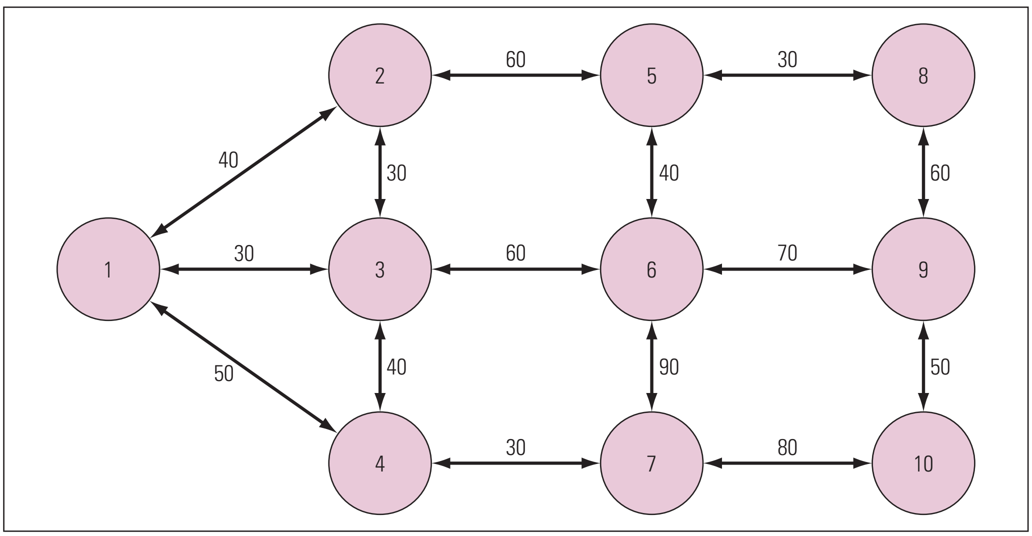
a. Create a Solver model that achieves Max's objective.
b. If Max used this model to predict the winner in each of last year's games, how many times would he have correctly picked the winning team?
c. Suppose Max wants to maximize the chance of picking the winning teams. Re-solve the problem to achieve this objective.
d. If Max used your model from part (c) to predict the winner in each of last year's games, how many times would he have correctly picked the winning team?

CASE 8-1 Tour de Europe

The summer before completing his MBA, Noah Franklin finally decided to take the trip to Europe that he had always dreamed about. However, given his limited bank account, he knows he will have to plan and budget wisely in order to go everywhere and see everything he wants to see. With some quick detective work on the Internet, Noah quickly found inexpensive sleeping quarters in each of the 10 cities he is interested in visiting. He also discovered there are several low-cost airlines providing no-frills transportation between various European cities. Figure 8.45 summarizes the possible airline flights between 10 different European cities with flight costs indicated on the arcs.

FIGURE 8.45

Possible flights for the Tour De Europe problem



Noah would really like to visit each of the 10 cities. Because his round trip flight from the United States arrives at and, following his vacation, departs from city 1, his tour of Europe needs to begin and end in city.

a. Construct a spreadsheet model Noah could use to determine the least costly way to visit all 10 European cities exactly once. What is the optimal itinerary for this problem and how much airfare would Noah have to pay?
b. Suppose the solution to the previous problem requires more money for airfare than Noah can afford. Construct a spreadsheet model Noah could use to determine the least costly way to visit cities 6, 8, and 10, starting and finishing from city 1. What is the optimal itinerary for this problem and how much airfare would Noah have to pay?

Electing the Next President

CASE 8-2

"So it's come down to this," thought Roger Mellichamp as he looked around at the empty Styrofoam coffee cups and papers littering his office. When he accepted the job of campaign manager for his long-time friend's run for the White House, he knew there would be long hours, lots of traveling, and constant media pressure. But the thing he most wanted to avoid was a close race with a final showdown just before the election. Roger knew that making decisions under those circumstances would be agonizing because the success of the campaign and, in many ways, the future of the country would hinge on those very decisions. Unfortunately, that's just where things stand.

With only two weeks before the U.S. presidential election, Roger's friend and the incumbent president are running neck-and-neck in the polls. So, Roger's plans for the final two weeks of the campaign will be critical, and he wants to make sure he uses the candidate's time and the campaign's remaining money in the most effective way. Although the outcome of the election has been pretty much decided in most states, the electoral votes from the states of Florida, Georgia, California, Texas, Illinois, New York, Virginia, and Michigan are still up for grabs by either candidate. Roger knows they must win as many of these states as possible if his friend is to become the next president.

Several weeks ago, it became evident that the race was going to be close. So, Roger hired a statistical consultant to estimate the percentage of votes the campaign will receive in each of the states based on the amount of money the campaign spends and the number of times the candidate visits each state during the final two weeks before the election. The results of the consultant's analysis provided the following function:

$$\text{Percentage of votes state } k = 1 - \text{EXP}(-\boldsymbol{a}\text{V}_k - \boldsymbol{b}\text{D}_k)$$

where

V_k = the number of times the candidate visits state k in the last two weeks of the campaign

D_k = the dollars (in \$1,000,000s) the campaign spends on advertising in state k in the last two weeks of the campaign

The following table summarizes the consultant's estimates of the parameters $\boldsymbol{a}$ and $\boldsymbol{b}$ for each state, along with the number of electoral votes at stake in each state:

State	*a*	*b*	Electoral Votes
Florida	0.085	0.31	25
Georgia	0.117	0.27	13
California	0.098	0.21	54
Texas	0.125	0.28	32
Illinois	0.128	0.26	22
New York	0.105	0.22	33
Virginia	0.134	0.24	13
Michigan	0.095	0.38	18

Roger believes the candidate can make 21 campaign stops in the next two weeks, and there is $15 million left in the campaign budget available for advertising. He wants to spend at least $500,000 in each of these states in the next two weeks. He also wants the candidate to make at least 1, but no more than 5, campaign stops in each of these states. Within these constraints, Roger wants to allocate these resources to maximize the number of electoral votes his candidate can receive. Assume a candidate needs 51% of the vote to win in each state.

a. Formulate an NLP model for this problem.
b. Implement your model in a spreadsheet and solve it.
c. How much money should Roger spend in each state?
d. How many campaign stops should the candidate make in each state?
e. What is the expected number of electoral votes generated by this solution?

CASE 8-3 Making Windows at Wella

Wella Corporation is a privately held manufacturer of doors and windows with annual sales in excess of $900 million. Some of the company's products are "standard size" and sold through wholesale and retail building material centers. However, much of their business involves manufacturing custom windows that can vary in size from 12 inches to 84 inches in quarter-inch increments.

The company has 2 plants, located in Iowa and Pennsylvania. Each plant has five production lines devoted to custom window manufacturing. Each of these production lines operates 8 hours a day, 5 days a week and produces 50 windows per hour.

Sash and frame parts for the windows are cut from standard size "stock" pieces of lumber 16 feet in length. These stock pieces are purchased from a supplier who takes various pieces of lumber of various lengths, cuts out the defects (knot holes, cracks, and so on), and finger joins the pieces together to create the 16-foot stock pieces that are basically free of defects.

Wella cuts all the sash and frame parts for a particular window and then immediately passes that set of parts to the next operation in the production process for further assembly (i.e., it does not carry inventories of parts of various length). However, the parts for any particular window may be cut in any order.

The demand for custom windows varies such that no two days (or even hours) of production are ever the same. Currently, line workers take a 16-foot piece of stock and start cutting parts for a window in the same order as they are listed on the bill of materials (BOM) until the remaining piece of stock is too short to cut the next required part.

As a simplified example, suppose the first window being produced has a BOM listing two 3-foot parts and two 4-foot parts (in that order). (Note that most of Wella's windows actually require eight or nine parts.) Those parts might be cut from a 16-foot stock piece and leave a 2-foot piece of scrap.

Now suppose the next window has a BOM that requires two 3-foot pieces and two 1-foot pieces (in that order). Because the 3-foot pieces can't be cut from the 2-foot scrap leftover from the first piece of stock, Wella would start cutting a new piece of stock. It seems to make more sense to use the 2-foot piece of scrap from the first piece of stock to cut the two 1-foot pieces required by the second window. However, reordering the pieces for the second window to eliminate the 2-foot piece of scrap could actually lead to the creation of a 3-foot piece of scrap later in the production process. (Any pieces of "scrap" at the end of one stock piece that cannot be used in the next job are indeed scrapped as moving these scrap pieces into and out of inventory becomes a logistical nightmare.)

Being able to “look ahead” to see the downstream impacts of reordering decisions is beyond the capability of most humans—especially when this must be repeated over and over on an ongoing basis. As a result, Wella wants to develop a system to optimize the production on each line on an hour-by-hour basis.

The file WellaData.xlsx contains data for a half-hour of production on one of Wella’s lines. (This line produces windows that require eight sash/frame parts per window.) Assume Wella wants to produce the windows in the order indicated, but the parts for each window can be produced in any order. Wella wants to determine an optimal part-cutting sequence that would allow the company to minimize the amount of scrap (and the number of stock pieces required to fill all the orders).

a. How many possible solutions are there to this problem?
b. Design a spreadsheet model for this problem. How many pieces of stock would have to be cut to produce the windows in this half-hour of production if Wella processes the windows and parts in the order given (in WellaData.xlsx)? How much scrap is generated by this solution?
c. Use Solver to optimize the problem. How many pieces of stock would have to be cut to produce the windows in this half-hour of production if Wella processes the windows and parts in the order Solver identifies? How much scrap is generated by this solution?
d. Assume that Wella pays $4 for each 16-foot piece of stock. If the results identified in the previous question are representative of the results that could be obtained on all of Wella’s production lines, how much money could Wella save over the course of a year?
e. What other suggestions/issues (if any) do you think Wella should consider before implementing your solution on their factory floor?

Newspaper Advertising Insert Scheduling

CASE 8-4

Advertising is the primary source of revenue for newspaper companies. Over the past 10 to15 years, the newspaper industry has been adjusting to changes in the mix of services that produce this revenue. The two main categories of services are run on press (ROP) advertising and preprinted insert advertising. ROP advertising is printed in the newspaper during the live press run each night, whereas preprinted inserts are produced (usually at a commercial printing facility) before the nightly production run and inserted into or delivered with the newspaper. Preprinted inserts offer several advantages for advertisers. Different sizes and quality of paper stock can be used to make ads unique and more colorful than possible on a newspaper printing press. Also, advertisers can tightly control quality for preprinted inserts unlike newspaper quality that varies widely.

Although revenue has been increasing in both categories of advertising, revenue from preprinted inserts has been growing at a higher rate than ROP advertising. For many newspaper companies, this shift in revenue mix has created scheduling challenges in the production area. With inserts, advertisers can select the zones to which specific sets of advertisements are distributed. A ***zone*** is a distinct geographical area where all the papers delivered in the area receive the same set of advertising inserts. The challenge for newspaper companies is to schedule the production run to process the correct combination of inserts for all the different zones and complete the run early enough to get the papers to the circulation department (and readers) on time. For many papers, the problem is exacerbated by advertisers’ desires to have “micro-zone” or more zones of smaller size, increasing the specificity with which different groups of consumers can be targeted.

Art Carter is the production manager for a medium-sized newspaper company. Each night, he and his employees must design a schedule for combining the appropriate advertising inserts for 36 different delivery zones into the newspaper. Art's company owns four inserting machines that can each be loaded with the inserts for a particular zone. Two of the inserting machines operate at a rate of 12,000 papers per hour, whereas the other two machines operate at a rate of 11,000 per hour. The equipment inserts the loaded set of inserts into newspapers coming off the production press until all the papers for a particular zone are completed.

When the inserts for a particular zone are completed, the inserting machine is stopped and reloaded with the inserts for the next zone. This reloading (or changeover) process takes different amounts of time depending on how much work is required to load the machine with the next zone's set of inserts. The zones can be processed in any order and on any of the four inserting machines. However, all the advertising for a particular zone must be processed on the same inserting machine (i.e., the inserts for a single zone are not distributed across multiple inserting machines).

The file NewspaperData.xlsx that accompanies this book contains sample data for a typical night's inserting workload for this company. In particular, this file contains the quantity of newspapers being produced for each of the 36 delivery zones and the estimated changeover times required to switch from one zone's set of inserts to another zone. Art has asked you to develop a model to design an optimal production schedule for the inserting equipment. In particular, he wants to determine which zones should be assigned to each of the four machines and the optimal order for processing the jobs assigned to each machine. His objective is to minimize the amount of time it takes (start to finish) to complete all the newspapers.