Wayne L. Winston

# Marketing Analytics

## Data-Driven Techniques with Microsoft Excel

# Marketing Analytics

## Data-Driven Techniques with Microsoft® Excel®

Wayne L. Winston

WILEY

*To my wonderful family: Gregory, Jennifer, and Vivian*

# Credits

**Executive Editor**
Robert Elliott

**Project Editor**
Victoria Swider

**Technical Editor**
Lopo Rego

**Production Editor**
Daniel Scribner

**Copy Editor**
San Dee Phillips

**Editorial Manager**
Mary Beth Wakefield

**Freelancer Editorial Manager**
Rosemarie Graham

**Associate Director of Marketing**
David Mayhew

**Marketing Manager**
Ashley Zurcher

**Business Manager**
Amy Knies

**Vice President and Executive Group Publisher**
Richard Swadley

**Associate Publisher**
Jim Minatel

**Project Coordinator, Cover**
Katie Crocker

**Proofreaders**
Josh Chase, Word One
Louise Watson, Word One

**Indexer**
Ron Strauss

**Cover Image**
Wiley

**Cover Designer**
Ryan Sneed

# About the Author

**Wayne Winston** is Professor Emeritus at the Indiana University Kelley School of Business and is currently a Visiting Professor at the University of Houston Bauer College of Business. Wayne has won more than 45 teaching awards at Indiana University. These awards include six school-wide MBA teaching awards. He has authored 25 reference journal articles and a dozen books including, *Operations Research*: *Applications and Algorithms* (Cengage, 1987), *Practical Management Science* (Cengage, 2011), *Data Analysis and Decision-Making* (Cengage, 2013), *Simulation Modeling with @RISK* (Cengage, 2004), *Mathletics* (Princeton, 2009), and *Excel 2013 Data Analysis and Business Modeling* (O'Reilly, 2014). Wayne has also developed two online courses for Harvard Business School: Spreadsheet Modeling, and Mathematics for Management. He has taught Excel modeling and consulted for many organizations including the U.S. Army, the U.S. Navy, Broadcom, Cisco, Intel, Pfizer, Eli Lilly, Ford, GM, PWC, Microsoft, IAC, Deloitte Consulting, Booz Allen Hamilton, QAS, eBay, the Dallas Mavericks, and the New York Knicks. Lastly, Wayne is a two-time *Jeopardy*! champion.

# About the Technical Editor

**Lopo Rego** joined the Kelley School of Business at Indiana University in 2011 as an Associate Professor of Marketing. Trained in Economics, he "converted to the dark side" during his MBA and has since been interested in understanding the association between marketing strategy and firm performance. This proved to be a life-long quest, leading him to Ann Arbor where he eventually earned his Ph.D. in Marketing at the University of Michigan's Ross School of Business. Not surprisingly, his research interests focus primarily in understanding how marketing decisions, strategies, and investments translate into firm performance, be it at the product-marketplace level, financial-accounting level or shareholder wealth level. Additionally, Lopo is interested in marketing analytics, namely in developing and analyzing marketing metrics that drive firm performance. His research has been published in such outlets as the *Journal of Marketing*, *Marketing Science*, *European Journal of Marketing*, *Journal of Empirical Generalisations in Marketing*, *Harvard Business Review*, *Journal of Research in Marketing,* and *Marketing Science Institute Working Paper Series*.

# Acknowledgments

**O**f all my books, this one was probably the hardest to write. Thanks to my wonderful wife Vivian who was so nice to me when I got frustrated during the authoring process. Wiley acquisitions editor Robert Elliott was always encouraging and his input was a great help in shaping the final product. Wiley project editor Victoria Swider did a great job in pushing me to become a better writer.

Lastly, I must give a special note of thanks to my technical editor, Associate Professor of Marketing at the Kelly School of Business, Lopo Rego. Lopo did an amazing job of suggesting alternative wording and catching errors. He went above and beyond his role as technical editor, and I am truly indebted to him for his Herculean efforts.

# Contents

# Contents

# Introduction

In the last 20 years, the use of analytic techniques in marketing has greatly increased. In April 2013, *Forbes* magazine reported a 67-percent growth in marketing-related analytics hires during the previous year and an amazing 136-percent growth during the previous 3 years.

Given this growth of interest in marketing analytics and my love of Excel modeling, I decided in 2004 to create a 7-week MBA elective in marketing analytics (K509) at the Indiana University Kelley School of Business. Although there are several excellent advanced marketing analytics books. (I am partial to *Database Marketing* by Robert Blattberg, Byung-Do Kim, and Scott Neslin (Springer, 2008).) I could not find an Excel-based book that provided a how-to-do-it approach suitable for an MBA elective or an advanced undergraduate course. With no suitable book in hand, I wrote up course notes that I used in classes for 10 years. The course has been wildly successful with nearly 65 percent of all MBA's at the Kelley School taking the class. In May 2013, I was honored to receive the Eli Lilly MBA teaching award as the best teacher in the MBA program, primarily for teaching K509. In November 2011, Robert Elliott of Wiley Publishing approached me about turning my notes into a book, and this book is the result. In addition to being utilized in K509, portions of the book have been used to teach marketing analytics to senior managers at Deloitte consulting, Booz Allen Hamilton consulting, and 3M marketing analysts.

## How This Book Is Organized

Since I started using Excel in classes in 1992, I have become a total convert to teaching by example. This book is no exception. Virtually every chapter's primary focus is to teach you the concepts through how-to examples. Each example has the following components:

- Step-by-step instructions
- A downloadable Excel file containing data and solutions
- Screenshots of various steps and sections of the Excel file for clarity

The downloadable Excel files provide complete solutions to the examples, but the instructions encourage you to follow along and work through them on your own. If you follow along using the provided Excel files, you can work in empty

cells alongside the completed solution and compare your result with the provided solution to ensure your success.

The book has been organized around 11 topical areas.

# Part I: Using Excel to Summarize Marketing Data

This part of the book introduces the marketing analyst to many Excel tools that can be used to analyze marketing problems: PivotTables (Chapter 1), charting (Chapter 2), and Excel statistical functions (Chapter 3), including the incredibly useful COUNTIF, COUNTIFS, SUMIF, SUMIFS, AVERAGEIF, and AVERAGEIFS functions.

# Part II: Pricing

The determination of a profit maximizing pricing strategy is always difficult. In this section you learn how to quickly estimate demand curves and use the Excel Solver (Chapter 4) to determine profit maximizing prices. The Excel Solver is then used to optimize price bundling (Chapter 5), nonlinear pricing strategies (Chapter 6), and price-skimming strategies (Chapter 7). A brief introduction to revenue management, also known as yield management (Chapter 8), is also included.

# Part III: Forecasting

Businesses need accurate forecasts of future sales. Sales forecasts drive decisions involving production schedules, inventory management, manpower planning, and many other parts of the business. In this section you first learn about two of the most used forecasting tools: simple linear (Chapter 9) and simple multiple regression (Chapters 10 and 11). Then you learn how to estimate the trend and seasonal aspects of sales (Chapter 12) and generate forecasts using two common extrapolation forecasting methods: the Ratio to Moving Average method (Chapter 13), and Winter's Method for exponential smoothing (Chapter 14) with trend and seasonality. Then you learn about neural networks (Chapter 15), a form of artificial intelligence whose role in marketing forecasting is rapidly growing.

# Part IV: What Do Customers Want?

Every brand manager wants to know how various product attributes drive the sales of a product. For example, what is most important in a consumer's choice of car: price, brand, engine horsepower, styling, or fuel economy? In this section you learn how conjoint analysis (Chapter 16) and discrete choice (Chapter 18) can be used to rank the importance of product attributes and also rank levels of product attributes. For example, what type of styling on an SUV is most preferred? You also

learn about the widely used tool of logistic regression (Chapter 17), which is used to estimate probabilities involving situations in which two, or binary, outcomes must be forecasted. For example, how a person's demographic information can be used to predict the chance that he will subscribe to a magazine.

## Part V: Customer Value

Companies cannot make intelligent decisions on how to spend money acquiring customers unless they understand the value of their customers. After all, spending $400 to acquire a customer who will generate $300 in long-term profits is a sure recipe for going out of business. In this section you learn how to measure customer value (Chapter 19), value companies based on the customer value concept (Chapter 20), incorporate uncertainty in customer value models (Chapter 21), and use your understanding of customer value to optimally allocate resources (Chapter 22) between acquisition and retention of customers.

## Part VI: Market Segmentation

No matter what product you sell, your market consists of different market segments. For example, in Chapter 23 you will use cluster analysis to show that that every U.S. city can be classified into one of four demographic segments. In Chapter 25 you will learn how classification trees can be used to segment a market. You are also introduced to the exciting concepts behind collaborative filtering (Chapter 24), which is the basis for Amazon.com and Netflix recommendations.

## Part VII: Forecasting New Product Sales

With little or no history about sales of a product, it is difficult to predict future product sales. Given a few data points, S curves (Chapter 26) can be used to predict future product sales. The famous Bass diffusion model (Chapter 27) explains how sales of products evolve over time and can be used to predict product sales even before a product comes to the market. The little-known Copernican Principle (Chapter 28) enables you to predict the remaining time for which a product will be sold.

## Part VIII: Retailing

Analytic techniques can help retailers deal with many important issues. The concepts of market basket analysis and lift (Chapter 29) help retailers derive a store layout that maximizes sales from complementary products. Recency, frequency, and monetary value analysis (Chapter 30) helps direct mailers maximize profit from their mailings. The widely known SCAN*PRO (Chapter 31) model helps retailers

determine how factors such as seasonality, price, and promotions influence product sales. In Chapter 32 you learn how to use analytic techniques to determine optimal allocation of store space between products and also optimize the use of a corporate sales force. Finally in Chapter 33 you learn how to forecast total sales of a product from a few data points.

## Part IX: Advertising

Department store owner John Wanamaker said, "Half the money I spend on advertising is wasted; the trouble is I don't know which half." In Chapter 34 you learn how John Wanamaker could have used the ADSTOCK model to measure the effectiveness of his advertising expenditures. In Chapter 35 you learn how to allocate ads between the available media outlets to maximize the effectiveness of ads. Chapter 36 deals with the math behind online ad auctions.

## Part X: Marketing Research Tools

Often the marketing analyst must deal with data sets involving many variables. Principal components (Chapter 37) and Multidimensional Scaling (Chapter 38) enable the marketing analysts to reduce data sets involving many variables to a few easily understood variables. Often the marketing analyst must classify objects into one of several groups. Naive Bayes and discriminant analysis (Chapter 39) are great tools for developing classification rules. When the marketing analyst wants to determine if a single factor or a pair of factors has a significant effect on product sales, ANOVA (Chapter 40 and Chapter 41) is a useful tool.

## Part XI: The Internet and Social Marketing

In the last 20 years, the Internet has turned our world upside down, and marketing is no exception. Social media such as Facebook and Twitter create many interesting opportunities for the marketer, which require careful analysis. In Chapter 42 you learn how the theory of networks sheds light on how you can identify people who are the key to spreading the word about your product. Chapter 43 discusses the math behind Malcom Gladwell's bestselling book *The Tipping Point (Back Bay Books, 2002)*. Chapter 44 discusses the math behind videos (such as the notorious "Gangnam Style") going viral. Finally, in Chapter 45 you learn how text mining can be used to glean useful insight from Twitter, blogs, and Facebook posts.

# Who Should Read This Book

There is plenty of material in this book for a one-semester course on marketing analytics at the advanced undergraduate or MBA level. I also believe the book can be useful to any corporate marketing analyst. With regard to prerequisites for the book, I assume you understand the Copy command in Excel. That is, you know when and where to put dollars signs in a formula. If you work hard, that's about all the prior knowledge needed to get a lot out of the book.

I always try to write my books in a modular fashion, so you can skip around and read about what interests you. If you don't want to read the book from start to finish, the following table should help you navigate the book.

| Chapters | Chapter Prerequisites |
|---|---|
| Chapter 1: Slicing and Dicing Marketing Data with PivotTables | None |
| Chapter 2: Using Excel Charts to Summarize Marketing Data | 1 |
| Chapter 3: Using Excel Functions to Summarize Marketing Data | 2 |
| Chapter 4: Estimating Demand Curves and Using Solver to Optimize Price | None |
| Chapter 5: Price Bundling | 4 |
| Chapter 6: Nonlinear Pricing | 5 |
| Chapter 7: Price Skimming and Sales | 5 |
| Chapter 8: Revenue Management | 4 |
| Chapter 9: Simple Linear Regression and Correlation | 3 |
| Chapter 10: Using Multiple Regression to Forecast Sales | 9 |
| Chapter 11: Forecasting in the Presence of Special Events | 10 |
| Chapter 12: Modeling Trend and Seasonality | 5 and 11 |
| Chapter 13: Ratio to Moving Average Forecasting Method | 3 and 12 |
| Chapter 14: Winter's Method | 12 |
| Chapter 15: Using Neural Networks to Forecast Sales | 10 |
| Chapter 16: Conjoint Analysis | 10 |

*Continues*

*(continued)*

| Chapters | Chapter Prerequisites |
| --- | --- |
| Chapter 17: Logistic Regression | 16 |
| Chapter 18: Discrete Choice Analysis | 17 |
| Chapter 19: Calculating Lifetime Customer Value | 3 |
| Chapter 20: Using Customer Value to Value a Business | 19 |
| Chapter 21: Customer Value, Monte Carlo Simulation, and Marketing Decision Making | 19 |
| Chapter 22: Allocating Marketing Resources between Customer Acquisition and Retention | 4 and 19 |
| Chapter 23: Cluster Analysis | 5 |
| Chapter 24: Collaborative Filtering | 23 |
| Chapter 25: Using Classification Trees for Segmentation | 24 |
| Chapter 26: Using S Curves to Forecast Sales of a New Product | 5 and 12 |
| Chapter 27: The Bass Diffusion Model | 26 |
| Chapter 28: Using the Copernican Principle to Predict Duration of Future Sales | None |
| Chapter 29: Market Basket Analysis and Lift | 19 |
| Chapter 30: RFM Analysis and Optimizing Direct Mail Campaigns | 29 |
| Chapter 31: Using the SCAN*PRO Model and Its Variants | 12 |
| Chapter 32: Allocating Retail Space and Sales Resources | 5 |
| Chapter 33: Forecasting Sales from Few Data Points | 31 |
| Chapter 34: Measuring the Effectiveness of Advertising | 31 |
| Chapter 35: Media Selection Models | 4, 21, and 34 |
| Chapter 36: Pay Per Click (PPC) Online Advertising | None |
| Chapter 37: Principal Component Analysis (PCA) | 10 and 23 |
| Chapter 38: Multidimensional Scaling (MDS) | 37 |
| Chapter 39: Classification Algorithms: Naive Bayes Classifier and Discriminant Analysis | 37 and 38 |

| Chapters | Chapter Prerequisites |
|---|---|
| Chapter 40: Analysis of Variance: One-way ANOVA | 10 |
| Chapter 41: Analysis of Variance: Two-way ANOVA | 40 |
| Chapter 42: Networks | None |
| Chapter 43: The Mathematics Behind *The Tipping Point* | 42 |
| Chapter 44: Viral Marketing | 10, 15, and 39 |
| Chapter 45: Text Mining | 3 |

For example, before reading Chapter 5 you need to have read Chapter 4; before reading Chapter 34 you need to have read Chapter 31, and so on.

## Tools You Need

To work through the vast majority of the book, all you need is Excel 2007, 2010, or 2013. Chapters 15, 21, and 35 require use of the Palisade.com Decision Tools Suite. You can download a 15-day trial version of the suite from `www.Palisade.com`.

## What's on the Website

From the book website (`www.wiley.com/go/marketinganalytics`) you can download all Excel files used in the book as well as answers to all of the Exercises at the end of each chapter.

## Errata

We make every effort to ensure that there are no errors in the text or in the code. However, no one is perfect, and mistakes do occur. If you find an error in this book, like a spelling mistake or a calculation error, we would be very grateful for your feedback. By sending in errata you may save another reader hours of frustration and at the same time you will be helping us provide even higher quality information.

To submit errata for this book go to `http://support.wiley.com` and complete the form on the Ask a Question tab there to send us the error you have found. We'll check

the information and, if appropriate, post a message to the book's errata page and fix the problem in subsequent editions of the book.

## Summary

A famous Chinese proverb (popularized by the late management guru Stephen Covey) states, "If you give a man a fish you feed him for a day. If you teach a man to fish you feed him for a lifetime." Hopefully this book will teach you enough about marketing analytics so you will be well equipped to develop your own quantitative marketing models for most problems that come your way. Happy modeling!

# Using Excel to Summarize Marketing Data

# 1 Slicing and Dicing Marketing Data with PivotTables

In many marketing situations you need to analyze, or "slice and dice," your data to gain important marketing insights. Excel PivotTables enable you to quickly summarize and describe your data in many different ways. In this chapter you learn how to use PivotTables to perform the following:

- Examine sales volume and percentage by store, month and product type.
- Analyze the influence of weekday, seasonality, and the overall trend on sales at your favorite bakery.
- Investigate the effect of marketing promotions on sales at your favorite bakery.
- Determine the influence that demographics such as age, income, gender and geographic location have on the likelihood that a person will subscribe to *ESPN: The Magazine.*

## Analyzing Sales at True Colors Hardware

To start analyzing sales you first need some data to work with. The `data` worksheet from the `PARETO.xlsx` file (available for download on the companion website) contains sales data from two local hardware stores (uptown store owned by Billy Joel and downtown store owned by Petula Clark). Each store sells 10 types of tape, 10 types of adhesive, and 10 types of safety equipment. Figure 1-1 shows a sample of this data.

Throughout this section you will learn to analyze this data using Excel PivotTables to answer the following questions:

- What percentage of sales occurs at each store?
- What percentage of sales occurs during each month?
- How much revenue does each product generate?
- Which products generate 80 percent of the revenue?

| | Y | Z | AA | AB |
|---|---|---|---|---|
| 7 | Product | Month | Store | Price |
| 8 | Tape 10 | April | downtown | $2.50 |
| 9 | Safety 8 | August | uptown | $10.00 |
| 10 | Safety 2 | February | uptown | $10.00 |
| 11 | Safety 8 | November | uptown | $10.00 |
| 12 | Tape 10 | October | uptown | $2.50 |
| 13 | Safety 8 | January | uptown | $10.00 |
| 14 | Safety 8 | December | downtown | $10.00 |
| 15 | Safety 1 | September | downtown | $12.00 |
| 16 | Safety 2 | May | uptown | $10.00 |
| 17 | Adhesive | July | uptown | $7.00 |
| 18 | Adhesive | March | uptown | $7.00 |
| 19 | Safety 8 | August | downtown | $10.00 |
| 20 | Safety 8 | October | downtown | $10.00 |
| 21 | Tape 10 | July | downtown | $2.50 |
| 22 | Safety 2 | February | downtown | $10.00 |

**Figure 1-1:** Hardware store data

## Calculating the Percentage of Sales at Each Store

The first step in creating a PivotTable is ensuring you have headings in the first row of your data. Notice that Row 7 of the example data in the data worksheet has the headings Product, Month, Store, and Price. Because these are in place, you can begin creating your PivotTable. To do so, perform the following steps:

1. Place your cursor anywhere in the data cells on the data worksheet, and then click PivotTable in the Tables group on the Insert tab. Excel opens the Create PivotTable dialog box, as shown in Figure 1-2, and correctly guesses that the data is included in the range Y7:AB1333.
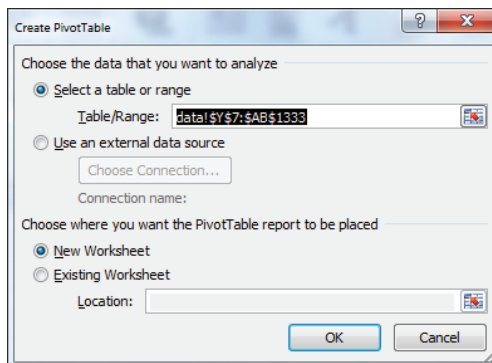


**Figure 1-2:** PivotTable Dialog Box

**NOTE**    If you select Use an External Data Source here, you could also refer to a database as a source for a PivotTable. In Exercise 14 at the end of the chapter you can practice creating PivotTables from data in different worksheets or even different workbooks.

2. Click OK and you see the PivotTable Field List, as shown in Figure 1-3.
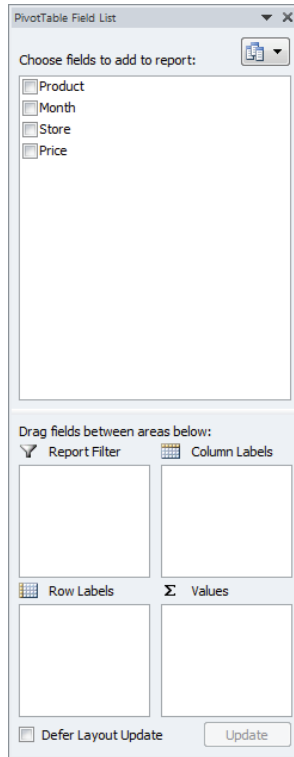


**Figure 1-3:** PivotTable Field List

3. Fill in the PivotTable Field List by dragging the PivotTable headings or *fields* into the boxes or *zones*. You can choose from the following four zones:

   ■ **Row Labels:** Fields dragged here are listed on the left side of the table in the order in which they are added to the box. In the current example, the Store field should be dragged to the Row Labels box so that data can be summarized by store.

- ■ **Column Labels:** Fields dragged here have their values listed across the top row of the PivotTable. In the current example no fields exist in the Column Labels zone.
- ■ **Values:** Fields dragged here are summarized mathematically in the PivotTable. The Price field should be dragged to this zone. Excel tries to guess the type of calculation you want to perform on a field. In this example Excel guesses that you want all Prices to be summed. Because you want to compute total revenue, this is correct. If you want to change the method of calculation for a data field to an average, a count, or something else, simply double-click the data field or choose Value Field Settings. You learn how to use the Value Fields Setting command later in this section.
- ■ **Report Filter:** Beginning in Excel 2007, *Report Filter* is the new name for the Page Field area. For fields dragged to the Report Filter zone, you can easily pick any subset of the field values so that the PivotTable shows calculations based only on that subset. In Excel 2010 or Excel 2013 you can use the exciting *Slicers* to select the subset of fields used in PivotTable calculations. The use of the Report Filter and Slicers is shown in the "Report Filter and Slicers" section of this chapter.

**NOTE** To see the field list, you need to be in a field in the PivotTable. If you do not see the field list, right-click any cell in the PivotTable, and select Show Field List.

Figure 1-4 shows the completed PivotTable Field List and the resulting PivotTable is shown in Figure 1-5 as well as on the `FirstorePT` worksheet.

Figure 1-5 shows the downtown store sold $4,985.50 worth of goods, and the uptown store sold $4,606.50 of goods. The total sales are $9592.

If you want a percentage breakdown of the sales by store, you need to change the way Excel displays data in the Values zone. To do this, perform these steps:

1. Right-click in the summarized data in the `FirstStorePT` worksheet and select Value Field Settings.
2. Select Show Values As and click the drop-down arrow on the right side of the dialog box.
3. Select the % of Column Total option, as shown in Figure 1-6.
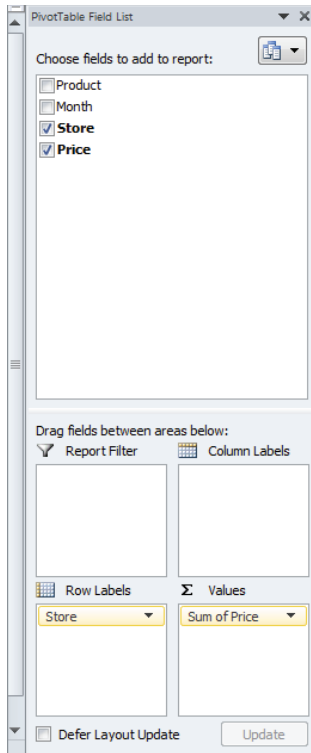
**Figure 1-4:** Completed PivotTable Field List



**Figure 1-5:** Completed PivotTable
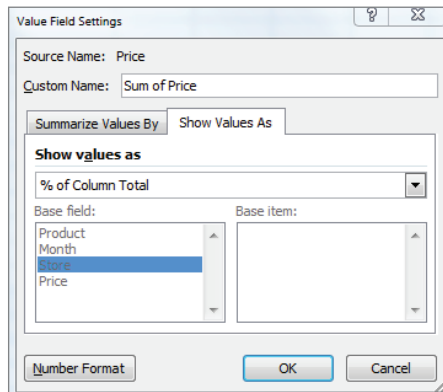


**Figure 1-6:** Obtaining percentage breakdown by Store

Figure 1-7 shows the resulting PivotTable with the new percentage breakdown by Store with 52 percent of the sales in the downtown store and 48 percent in the uptown store. You can also see this in the `revenue by store` worksheet of the `PARETO.xlsx` file.

**NOTE**   If you want a PivotTable to incorporate a different set of data, then under Options, you can select Change Data Source and select the new source data. To have a PivotTable incorporate changes in the original source data, simply right-click and select Refresh. If you are going to add new data below the original data and you want the PivotTable to include the new data when you select Refresh, you should use the Excel Table feature discussed in Chapter 2, "Using Excel Charts to Summarize Marketing Data."

| | A | B |
|---|---|---|
| 1 | | |
| 2 | | |
| 3 | Row Labels ▾ | Sum of Price |
| 4 | downtown | 51.98% |
| 5 | uptown | 48.02% |
| 6 | Grand Total | 100.00% |

**Figure 1-7:** Percentage breakdown by Store

## Summarizing Revenue by Month

You can also use a PivotTable to break down the total revenue by month and calculate the percentage of sales that occur during each month. To accomplish this, perform the following steps:

1. Return to the `data` worksheet and bring up the PivotTable Field List by choosing Insert PivotTable.
2. Drag the Month field to the Row Labels zone and the Price field to the Values zone. This gives the total sales by month. Because you also want a percentage breakdown of sales by month, drag the Price field again to the Values zone.
3. As shown in Figure 1-8, right-click on the first column in the Values zone and choose Value Field Settings; then choose the % of Column Total option. You now see the percentage monthly breakdown of revenue.
4. Double-click the Column headings and change them to Percentage of Sales by Month and Total Revenue.

5. Finally, double-click again the Total Revenue Column; select Number Format, and choose the Currency option so the revenue is formatted in dollars.

| | A | B | C |
|---|---|---|---|
| 1 | | | |
| 2 | | | |
| 3 | Row Labels | Percentage of Sales by Month | Total Revenue |
| 4 | January | 8.81% | $845.00 |
| 5 | February | 8.55% | $820.00 |
| 6 | March | 6.58% | $631.00 |
| 7 | April | 8.10% | $776.50 |
| 8 | May | 8.48% | $813.00 |
| 9 | June | 8.87% | $850.50 |
| 10 | July | 8.77% | $841.00 |
| 11 | August | 7.13% | $684.00 |
| 12 | September | 8.71% | $835.50 |
| 13 | October | 7.85% | $753.00 |
| 14 | November | 9.11% | $873.50 |
| 15 | December | 9.06% | $869.00 |
| 16 | Grand Total | 100.00% | $9,592.00 |

**Figure 1-8:** Monthly percentage breakdown of Revenue

You can see that $845 worth of goods was sold in January and 8.81 percent of the sales were in January. Because the percentage of sales in each month is approximately 1/12 (8.33 percent), the stores exhibit little seasonality. Part III, "Forecasting Sales of Existing Products," includes an extensive discussion of how to estimate seasonality and the importance of seasonality in marketing analytics.

## Calculating Revenue for Each Product

Another important part of analyzing data includes determining the revenue generated by each product. To determine this for the example data, perform the following steps:

1. Return to the `data` worksheet and drag the Product field to the Row Labels zone and the Price field to the Values zone.
2. Double-click on the Price column, change the name of the Price column to Revenue, and then reformat the Revenue Column as Currency.
3. Click the drop-down arrow in cell A3 and select Sort A to Z so you can alphabetize the product list and obtain the PivotTable in the `products` worksheet, as shown in Figure 1-9.

| | A | B |
|---|---|---|
| 3 | Row Labels ▼ | Revenue |
| 4 | Adhesive 1 | $24.00 |
| 5 | Adhesive 10 | $70.00 |
| 6 | Adhesive 2 | $63.00 |
| 7 | Adhesive 3 | $49.00 |
| 8 | Adhesive 4 | $238.00 |
| 9 | Adhesive 5 | $168.00 |
| 10 | Adhesive 6 | $42.00 |
| 11 | Adhesive 7 | $70.00 |
| 12 | Adhesive 8 | $42.00 |
| 13 | Adhesive 9 | $91.00 |
| 14 | Safety 1 | $732.00 |
| 15 | Safety 10 | $90.00 |
| 16 | Safety 2 | $740.00 |
| 17 | Safety 3 | $140.00 |
| 18 | Safety 4 | $90.00 |
| 19 | Safety 5 | $50.00 |
| 20 | Safety 6 | $220.00 |
| 21 | Safety 7 | $420.00 |
| 22 | Safety 8 | $5,050.00 |
| 23 | Safety 9 | $40.00 |
| 24 | Tape 1 | $93.00 |
| 25 | Tape 10 | $490.00 |
| 26 | Tape 2 | $25.00 |
| 27 | Tape 3 | $15.00 |
| 28 | Tape 4 | $35.00 |
| 29 | Tape 5 | $47.50 |
| 30 | Tape 6 | $67.50 |
| 31 | Tape 7 | $340.00 |
| 32 | Tape 8 | $20.00 |
| 33 | Tape 9 | $30.00 |
| 34 | Grand Total | $9,592.00 |

**Figure 1-9:** Sales by Product

You can now see the revenue that each product generated individually. For example, Adhesive 1 generated $24 worth of revenue.

## The Pareto 80–20 Principle

When slicing and dicing data you may encounter a situation in which you want to find which set of products generates a certain percentage of total sales. The well-known *Pareto 80–20 Principle* states that for most companies 20 percent of their products generate around 80 percent of their sales. Other examples of the Pareto Principle include the following:

- Twenty percent of the population has 80 percent of income.
- Of all possible problems customers can have, 20 percent of the problems cause 80 percent of all complaints.

To determine a percentage breakdown of sales by product, perform the following steps:

1. Begin with the PivotTable in the `products` worksheet and click the drop-down arrow in cell A3.

2. Select Value Filters; then choose Top 10…
3. Change the settings, as shown in Figure 1-10, to choose the products generating 80 percent of the revenue.
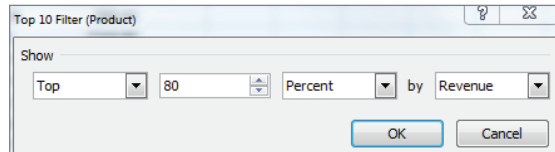


**Figure 1-10:** Using Value Filters to select products generating 80% of sales

The resulting PivotTable appears in the `Top 80%` worksheet (see Figure 1-11) and shows that the six products displayed in Figure 1-11 are the smallest set of products generating at least 80 percent of the revenue. Therefore only 20 percent of the products (6 out of 30) are needed to generate 80 percent of the sales.

**NOTE**    By clicking the funnel you may clear our filters, if desired.



| | A | B |
|---|---|---|
| 3 | Row Labels | Revenue |
| 4 | Safety 8 | $5,050.00 |
| 5 | Safety 2 | $740.00 |
| 6 | Safety 1 | $732.00 |
| 7 | Tape 10 | $490.00 |
| 8 | Safety 7 | $420.00 |
| 9 | Tape 7 | $340.00 |
| 10 | Grand Total | $7,772.00 |

**Figure 1-11:** 6 Products Generate 80% of Revenue

## The Report Filter and Slicers

One helpful tool for analyzing data is the Report Filter and the exciting Excel 2010 and 2013 Slicers Feature. Suppose you want to break down sales from the example data by month and store, but you feel showing the list of products in the Row or Column Labels zones would clutter the PivotTable. Instead, you can drag the Month field to the Row Labels zone, the Store field to the Column Labels zone, the Price field to the Value zone, and the Product field to the Report Filter zone. This yields the PivotTable in worksheet `Report filter unfiltered`, as shown in Figure 1-12.

| | A | B | C | D |
|---|---|---|---|---|
| 1 | Product | (All) | | |
| 2 | | | | |
| 3 | Sum of Price | Column Labels | | |
| 4 | Row Labels | downtown | uptown | Grand Total |
| 5 | January | $482.00 | $363.00 | $845.00 |
| 6 | February | $363.00 | $457.00 | $820.00 |
| 7 | March | $299.00 | $332.00 | $631.00 |
| 8 | April | $407.00 | $369.50 | $776.50 |
| 9 | May | $408.50 | $404.50 | $813.00 |
| 10 | June | $400.50 | $450.00 | $850.50 |
| 11 | July | $446.00 | $395.00 | $841.00 |
| 12 | August | $372.50 | $311.50 | $684.00 |
| 13 | September | $446.50 | $389.00 | $835.50 |
| 14 | October | $394.00 | $359.00 | $753.00 |
| 15 | November | $503.50 | $370.00 | $873.50 |
| 16 | December | $463.00 | $406.00 | $869.00 |
| 17 | Grand Total | 4985.5 | 4606.5 | 9592 |

**Figure 1-12:** PivotTable used to illustrate Slicers
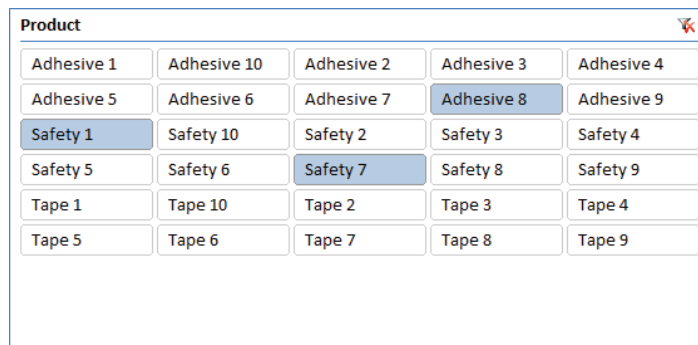
By clicking the drop-down arrow in the Report Filter, you can display the total revenue by Store and Month for any subset of products. For example, if you select products Safety 1, Safety 7, and Adhesive 8, you can obtain the PivotTable in the `Filtered with a slicer` worksheet, as shown in Figure 1-13. You see here that during May, sales of these products downtown are $10.00 and uptown are $34.00.

| | A | B | C | D |
|---|---|---|---|---|
| 1 | Product | (Multiple Items) | | |
| 2 | | | | |
| 3 | Sum of Price | Column Labels | | |
| 4 | Row Labels | downtown | uptown | Grand Total |
| 5 | January | $46.00 | $22.00 | $68.00 |
| 6 | February | $44.00 | $68.00 | $112.00 |
| 7 | March | $51.00 | $63.00 | $114.00 |
| 8 | April | $44.00 | $44.00 | $88.00 |
| 9 | May | $10.00 | $34.00 | $44.00 |
| 10 | June | $89.00 | $56.00 | $145.00 |
| 11 | July | $49.00 | | $49.00 |
| 12 | August | $10.00 | $36.00 | $46.00 |
| 13 | September | $88.00 | $44.00 | $132.00 |
| 14 | October | $102.00 | $87.00 | $189.00 |
| 15 | November | $56.00 | $70.00 | $126.00 |
| 16 | December | $32.00 | $49.00 | $81.00 |
| 17 | Grand Total | 621 | 573 | 1194 |

**Figure 1-13:** PivotTable showing sales for Safety 1, Safety 7 and Adhesive 8

As you can see from Figure 1-13, it is difficult to know which products were used in the PivotTable calculations. The new Slicer feature in Excel 2010 and 2013 (see the `Filtered with a slicer` worksheet) remedies this problem. To use this tool perform the following steps:

1. Put your cursor in the PivotTable in the `Filtered with a Slicer` worksheet and select Slicer from the Insert tab.
2. Select Products from the dialog box that appears and you see a Slicer that enables you to select any subset of products (select a product and then hold down the Control Key to select another product) from a single column.
3. Click inside the Slicer and you will see Slicer Tools on the ribbon. After selecting the Buttons section from Slicer Tools change Columns to 5. Now the products show up in five columns (see Figure 1-14).

| Product | | | | |
|---|---|---|---|---|
| Adhesive 1 | Adhesive 10 | Adhesive 2 | Adhesive 3 | Adhesive 4 |
| Adhesive 5 | Adhesive 6 | Adhesive 7 | Adhesive 8 | Adhesive 9 |
| Safety 1 | Safety 10 | Safety 2 | Safety 3 | Safety 4 |
| Safety 5 | Safety 6 | Safety 7 | Safety 8 | Safety 9 |
| Tape 1 | Tape 10 | Tape 2 | Tape 3 | Tape 4 |
| Tape 5 | Tape 6 | Tape 7 | Tape 8 | Tape 9 |

**Figure 1-14:**  Slicer selection for sales of Safety 1, Safety 7 and Adhesive 8.

A Slicer provides sort of a "dashboard" to filter on subsets of items drawn from a PivotTable field(s). The Slicer in the `Filtered with a slicer` worksheet makes it obvious that the calculations refer to Safety 1, Safety 7, and Adhesive 8. If you hold down the Control key, you can easily resize a Slicer.

**NOTE**    If you double-click in a cell in a PivotTable, Excel drills down to the source data used for that cell's calculations and places the source data in a separate sheet. For example, if in the `Report filtered unfiltered` worksheet you double-click in the January downtown cell, you can obtain the source data in the worksheet `January downtown`, as shown in Figure 1-15.

You have learned how PivotTables can be used to slice and dice sales data. Judicious use of the Value Fields Settings capability is often the key to performing the needed calculations.

| | A | B | C | D |
|---|---|---|---|---|
| 1 | Product | Month | Store | Price |
| 2 | Safety 7 | January | downtow | 10 |
| 3 | Safety 1 | January | downtow | 12 |
| 4 | Safety 1 | January | downtow | 12 |
| 5 | Safety 1 | January | downtow | 12 |

**Figure 1-15:** Drilling down on January Downtown Sales

# Analyzing Sales at La Petit Bakery

La Petit Bakery sells cakes, pies, cookies, coffee, and smoothies. The owner has hired you to analyze factors affecting sales of these products. With a PivotTable and all the analysis skills you now have, you can quickly describe the important factors affecting sales. This example paves the way for a more detailed analysis in Part III of this book.

The file `BakeryData.xlsx` contains the data for this example and the file `LaPetitBakery.xlsx` contains all the completed PivotTables. In the `Bakerydata .xlsx` workbook you can see the underlying daily sales data recorded for the years 2013–2015. Figure 1-16 shows a subset of this data.

| | A | B | C | D | E | F | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 | | | | | daywk | weekday | Date | Cakes | Pies | Cookies | Smoothies | Coffee | promotion |
| 5 | | | | | Tuesday | 2 | 1/1/2013 | 79 | 46 | 518 | 60 | 233 | none |
| 6 | | 1 Monday | | | Wednesd | 3 | 1/2/2013 | 91 | 50 | 539 | 161 | 427 | none |
| 7 | | 2 Tuesday | | | Thursday | 4 | 1/3/2013 | 47 | 60 | 222 | 166 | 347 | none |
| 8 | | 3 Wednesday | | | Friday | 5 | 1/4/2013 | 89 | 64 | 734 | 153 | 358 | none |
| 9 | | 4 Thursday | | | Saturday | 6 | 1/5/2013 | 112 | 73 | 764 | 240 | 392 | none |
| 10 | | 5 Friday | | | Sunday | 7 | 1/6/2013 | 89 | 57 | 922 | 259 | 510 | none |
| 11 | | 6 Saturday | | | Monday | 1 | 1/7/2013 | 70 | 50 | 476 | 120 | 334 | none |
| 12 | | 7 Sunday | | | Tuesday | 2 | 1/8/2013 | 70 | 48 | 496 | 222 | 316 | none |
| 13 | | | | | Wednesd | 3 | 1/9/2013 | 59 | 37 | 587 | 181 | 156 | none |
| 14 | | | | | Thursday | 4 | 1/10/2013 | 71 | 36 | 488 | 178 | 298 | none |
| 15 | | | | | Friday | 5 | 1/11/2013 | 74 | 50 | 645 | 100 | 490 | none |
| 16 | | | | | Saturday | 6 | 1/12/2013 | 119 | 71 | 438 | 162 | 416 | none |
| 17 | | | | | Sunday | 7 | 1/13/2013 | 90 | 51 | 568 | 137 | 434 | none |
| 18 | | | | | Monday | 1 | 1/14/2013 | 96 | 48 | 585 | 194 | 573 | promotion |

**Figure 1-16:** Data for La Petit Bakery

A 1 in the weekday column indicates the day was Monday, a 2 indicates Tuesday, and so on. You can obtain these days of the week by entering the formula = `WEEKDAY(G5,2)` in cell F5 and copying this formula from F5 to the range

F6:F1099. The second argument of 2 in the formula ensures that a Monday is recorded as 1, a Tuesday as 2, and so on. In cell E5 you can enter the formula =VLOOKUP(F5,lookday,2) to transform the 1 in the weekday column to the actual word Monday, the 2 to Tuesday, and so on. The second argument lookday in the formula refers to the cell range A6:B12.

> **NOTE**  To name this range lookday simply select the range and type lookday in the Name box (the box directly to the left of the Function Wizard) and press Enter. Naming a range ensures that Excel knows to use the range lookday in any function or formula containing lookday.

The VLOOKUP function finds the value in cell F5 (2) in the first column of the lookday range and replaces it with the value in the same row and second column of the lookday range (Tuesday.) Copying the formula =VLOOKUP(F5,lookday,2) from E5 to E6:E1099 gives you the day of the week for each observation. For example, on Friday, January 11, 2013, there was no promotion and 74 cakes, 50 pies, 645 cookies, 100 smoothies, and 490 cups of coffee were sold.

Now you will learn how to use PivotTables to summarize how La Petit Bakery's sales are affected by the following:

- Day of the week
- Month of the year
- An upward (or downward!) trend over time in sales
- Promotions such as price cuts

## Summarizing the Effect of the Day of the Week on Bakery Sales

La Petit Bakery wants to know how sales of their products vary with the day of the week. This will help them better plan production of their products.

In the day of week worksheet you can create a PivotTable that summarizes the average daily number of each product sold on each day of the week (see Figure 1-17). To create this PivotTable, perform the following steps:

1. Drag the daywk field to the Row Labels zone and drag each product to the Values zone.
2. Double-click each product, and change the summary measure to Average. You'll see, for example, that an average of 96.5 cakes was sold on Sunday.
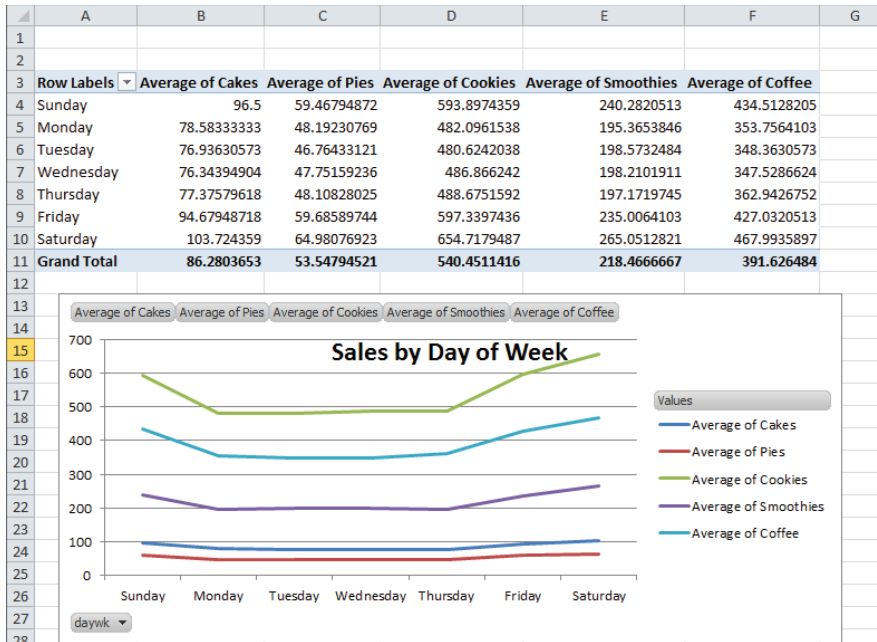
**Figure 1-17:** Daily breakdown of Product Sales

As the saying (originally attributed to Confucius) goes, "A picture is worth a thousand words." If you click in a PivotTable and go up to the Options tab and select PivotChart, you can choose any of Excel's chart options to summarize the data (Chapter 2, "Using Excel Charts to Summarize Marketing Data," discusses Excel charting further). Figure 1-17 (see the `Daily Breakdown` worksheet) shows the first Line option chart type. To change this, right-click any series in a chart. The example chart here shows that all products sell more on the weekend than during the week. In the lower left corner of the chart, you can filter to show data for any subset of weekdays you want.

## Analyzing Product Seasonality

If product sales are approximately the same during each month, they do not exhibit seasonality. If, however, product sales are noticeably higher (or lower) than average during certain quarters, the product exhibits seasonality. From a marketing standpoint, you must determine the presence and magnitude of seasonality to more efficiently plan advertising, promotions, and manufacturing decisions and investments. Some real-life illustrations of seasonality include the following:

■ Amazon's fourth quarter revenues are approximately 33 percent higher than an average quarter. This is because of a spike in sales during Christmas.

■ Tech companies such as Microsoft and Cisco invariably have higher sales during the last month of each quarter and reach maximum sales during the last month of the fiscal year. This is because the sales force doesn't get its bonuses unless it meets quarterly or end of year quotas.

To determine if La Petit Bakery products exhibit seasonality, you can perform the following steps:

1. Begin with your cursor anywhere in the data in the `BakeryData.xlsx` workbook. From the Insert tab select PivotTable and the PivotTable Field List will appear. Drag the Date field to the Row Labels zone and as before, drag each product to the Values zone and again change the entries in the Values zone to average sales for each product.

2. At first you see sales for every day, but you actually want to group sales by month. To do this, put the cursor on any date, right-click, and choose Group.

3. To group the daily sales into monthly buckets, choose Months from the dialog box, as shown in Figure 1-18.
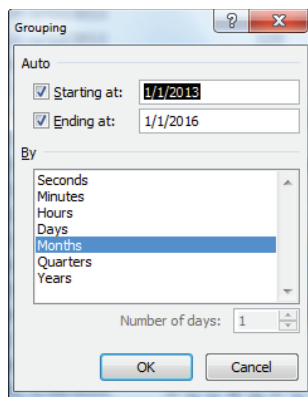


**Figure 1-18:** Grouping data by Month

4. Now select PivotChart from the Options tab. After selecting the first Line chart option you obtain the PivotChart and PivotTable, as shown in the `monthly breakdown` worksheet of the `LaPetitBakery.xlsx` file (see Figure 1-19).

| Row Labels ▼ | Average of Cakes | Average of Pies | Average of Cookies | Average of Smoothies | Average of Coffee |
|---|---|---|---|---|---|
| Jan | 83.07526882 | 51.64516129 | 516.8817204 | 171.5591398 | 398.2688172 |
| Feb | 84.39285714 | 53.6547619 | 534.3571429 | 174.9047619 | 398.9761905 |
| Mar | 84.38709677 | 53.75268817 | 545.8494624 | 182.0430108 | 390.5053763 |
| Apr | 84.45555556 | 51.63333333 | 519.3333333 | 211.2888889 | 383.8888889 |
| May | 89.89247312 | 52.10752688 | 536.1827957 | 241.2258065 | 369.8494624 |
| Jun | 86.22222222 | 54.97777778 | 525.5333333 | 277.4555556 | 388.2 |
| Jul | 86.93548387 | 52.48387097 | 540.0860215 | 286.6021505 | 398.0107527 |
| Aug | 84.92473118 | 55.16129032 | 544.8817204 | 287.3548387 | 406.2795699 |
| Sep | 87.13333333 | 52.35555556 | 551.1666667 | 219.3 | 378.6777778 |
| Oct | 86.46236559 | 55.02150538 | 554.0537634 | 213.0430108 | 392.6666667 |
| Nov | 89.38888889 | 55.05555556 | 573.5222222 | 171.8111111 | 396.8111111 |
| Dec | 87.97849462 | 54.7311828 | 543.2258065 | 180.9892473 | 397.483871 |
| Grand Total | 86.2803653 | 53.54794521 | 540.4511416 | 218.4666667 | 391.626484 |



**Figure 1-19:** Monthly breakdown of Bakery Sales

This chart makes it clear that smoothie sales spike upward in the summer, but sales of other products exhibit little seasonality. In Part III of this book you can find an extensive discussion of how to estimate seasonality.

Given the strong seasonality and corresponding uptick in smoothie sales, the bakery can probably "trim" advertising and promotions expenditures for smoothies between April and August. On the other hand, to match the increased demand for smoothies, the bakery may want to guarantee the availability and delivery of the ingredients needed for making its smoothies. Similarly, if the increased demand places stress on the bakery's capability to serve its customers, it may consider hiring extra workers during the summer months.

**NOTE**    If you right-click any month and select Ungroup, you can undo the grouping and return to a display of the daily sales data.

## Analyzing the Trend in Bakery Sales

The owners of La Petit Bakery want to know if sales are improving. Looking at a graph of each product's sales by month will not answer this question if seasonality is present. For example, Amazon.com has lower sales every January than the previous month due to Christmas. A better way to analyze this type of trend is to compute and chart average daily sales for each year. To perform this analysis, complete the following steps:

1. Put your cursor inside the data in the `Data` worksheet of the `BakeryData` `.xlsx` file, create a PivotTable, and drag each product to the Values zone and again change the method of summary from Sum to Average.

2. Then drag the Date field to the Row Labels zone, place the cursor on any date, and right-click Group and choose Years. You see a monthly summary of average daily sales for each Month and Year.

3. Drag the Date field away from the Row Labels zone and you are left with the summary of product sales by year, as shown in Figure 1-20 (see the work in worksheet `by Year`).

4. As you did before, create a Line PivotChart. The chart shows sales for each product are trending upward, which is good news for the client.

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 3 | Row Labels | Average of Cakes | Average of Pies | Average of Cookies | Average of Smoothies | Average of Coffee |
| 4 | 2013 | 81.84383562 | 52.65205479 | 523.2986301 | 213.8082192 | 377.1863014 |
| 5 | 2014 | 86.90136986 | 53.74520548 | 542.8136986 | 219.4684932 | 394.5479452 |
| 6 | 2015 | 90.09589041 | 54.24657534 | 555.2410959 | 222.1232877 | 403.1452055 |
| 7 | Grand Total | 86.2803653 | 53.54794521 | 540.4511416 | 218.4666667 | 391.626484 |

**Figure 1-20:** Summary of Sales by Year

The annual growth rates for products vary between 1.5 percent and 4.9 percent. Cake sales have grown at the fastest rate, but represent a small part of overall sales. Cookies and coffee sales have grown more slowly, but represent a much larger percentage of revenues.

## Analyzing the Effect of Promotions on Sales

To get a quick idea of how promotions affect sales, you can determine average sales for each product on the days you have promotions and compare the average on days with promotions to the days without promotions. To perform these computations keep the same fields in the Value zone as before, and drag the Promotion field to the Row Labels zone. After creating a line PivotChart, you can obtain the results, as shown in Figure 1-21 (see the `promotion` worksheet).
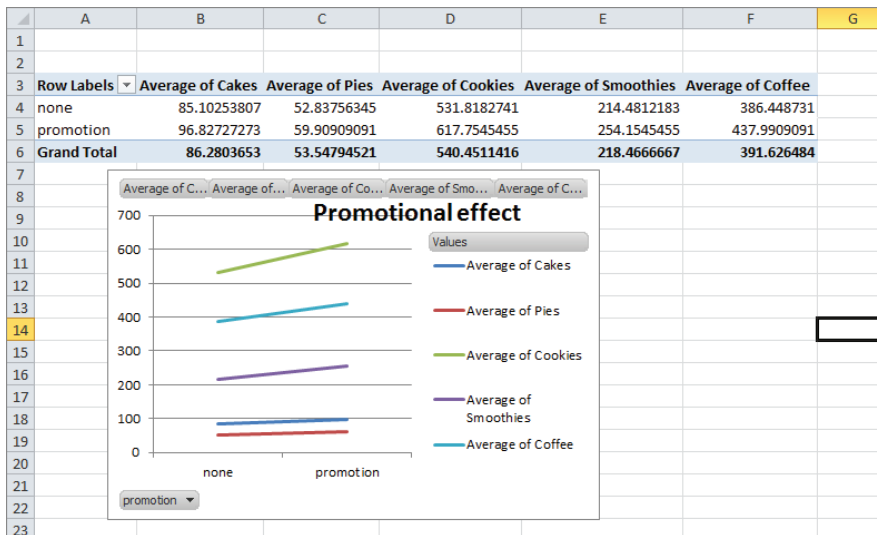


**Figure 1-21:** Effect of Promotion on Sales

The chart makes it clear that sales are higher when you have a promotion. Before concluding, however, that promotions increase sales, you must "adjust" for other factors that might affect sales. For example, if all smoothie promotions occurred during summer days, seasonality would make the average sales of smoothies on promotions higher than days without promotions, even if promotions had no real effect on sales. Additional considerations for using promotions include costs; if the costs of the promotions outweigh the benefits, then the promotion should not be undertaken. The

marketing analyst must be careful in computing the benefits of a promotion. If the promotion yields new customers, then the long-run value of the new customers (see Part V) should be included in the benefit of the promotion. In Parts VIII and IX of this book you can learn how to perform more rigorous analysis to determine how marketing decisions such as promotions, price changes, and advertising influence product sales.

# Analyzing How Demographics Affect Sales

Before the marketing analyst recommends where to advertise a product (see Part IX), she needs to understand the type of person who is likely to purchase the product. For example, a heavy metal fan magazine is unlikely to appeal to retirees, so advertising this product on a television show that appeals to retirees (such as *Golden Girls*) would be an inefficient use of the advertising budget. In this section you will learn how to use PivotTables to describe the demographic of people who purchase a product.

Take a look at the data worksheet in the espn.xlsx file. This has demographic information on a randomly chosen sample of 1,024 subscribers to *ESPN: The Magazine*. Figure 1-22 shows a sample of this data. For example, the first listed subscriber is a 72-year-old male living in a rural location with an annual family income of $72,000.

## Analyzing the Age of Subscribers

One of the most useful pieces of demographic information is age. To describe the age of subscribers, you can perform the following steps:

1. Create a PivotTable by dragging the Age field to the Row Labels zone and the Age field to the Values zone.
2. Unfortunately, Excel assumes that you want to calculate the Sum of Ages. Double-click the Sum of Ages heading, and change this to Count of Age.
3. Use Value Field settings with the % of Column Total setting to show a percentage breakdown by age.
4. Finally, right-click on any listed age in the PivotTable and select Group. This enables you to group ages in 10-year increments. You can also use the PivotChart feature using the first Column chart option to create a column

chart (see Figure 1-23 and the `age` worksheet) to summarize the age distribution of subscribers.

| | G | H | I | J |
|---|---|---|---|---|
| 3 | Age | Gender | Income | Location |
| 4 | 72 | m | 72 | rural |
| 5 | 29 | m | 68 | suburban |
| 6 | 33 | m | 57 | suburban |
| 7 | 25 | m | 62 | suburban |
| 8 | 38 | m | 164 | urban |
| 9 | 33 | m | 44 | urban |
| 10 | 18 | f | 62 | rural |
| 11 | 17 | m | 68 | urban |
| 12 | 32 | f | 53 | urban |
| 13 | 24 | f | 92 | urban |
| 14 | 26 | f | 54 | suburban |
| 15 | 40 | m | 88 | urban |
| 16 | 26 | f | 46 | rural |
| 17 | 21 | m | 83 | suburban |
| 18 | 29 | m | 144 | rural |
| 19 | 51 | f | 30 | suburban |
| 20 | 17 | m | 47 | urban |
| 21 | 29 | m | 61 | urban |
| 22 | 92 | m | 123 | urban |
| 23 | 20 | m | 76 | suburban |
| 24 | 20 | f | 57 | rural |
| 25 | 65 | m | 104 | urban |
| 26 | 26 | m | 46 | urban |
| 27 | 74 | m | 113 | urban |
| 28 | 37 | f | 109 | rural |

**Figure 1-22:** Demographic data for *ESPN: The Magazine*

You find that most of the magazine's subscribers are in the 18–37 age group. This knowledge can help ESPN find TV shows to advertise that target the right age demographic.

## Analyzing the Gender of Subscribers

You can also analyze the gender demographics of *ESPN: The Magazine* subscribers. This will help the analyst to efficiently spend available ad dollars. After all, if all subscribers are male, you probably do not want to place an ad on *Project Runway*.

1. In the `data` worksheet, drag the Gender field to the Column Labels zone and the Gender field to the Values zone.
2. Right-click the data and use Value Field Settings to change the calculations to Show Value As % of Row Total; this enables you to obtain the PivotTable shown in the `gender` worksheet (see Figure 1-24.)

| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 3 | Row Labels | Fraction of subscribers | | | | | |
| 4 | 8-17 | 8.20% | | | | | |
| 5 | 18-27 | 28.81% | | | | | |
| 6 | 28-37 | 28.81% | | | | | |
| 7 | 38-47 | 17.58% | | | | | |
| 8 | 48-57 | 8.01% | | | | | |
| 9 | 58-67 | 2.64% | | | | | |
| 10 | 68-77 | 3.22% | | | | | |
| 11 | 78-87 | 0.78% | | | | | |
| 12 | 88-97 | 1.17% | | | | | |
| 13 | 98-107 | 0.20% | | | | | |
| 14 | 108-117 | 0.59% | | | | | |
| 15 | Grand Total | 100.00% | | | | | |
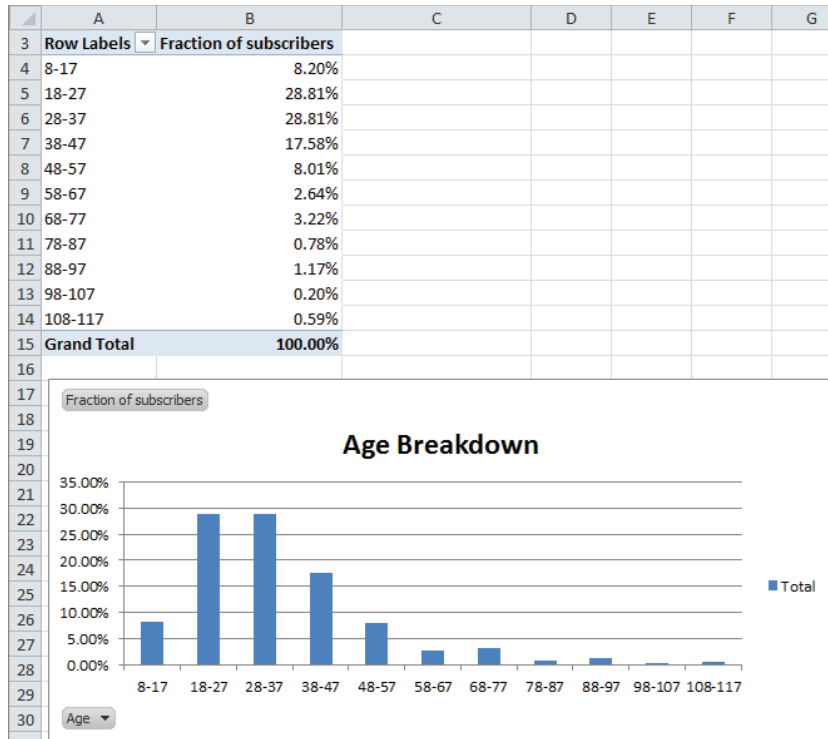
**Figure 1-23:** Age distribution of subscribers

| | A | B | C | D |
|---|---|---|---|---|
| 1 | | | | |
| 2 | | | | |
| 3 | | Column Labels | | |
| 4 | | f | m | Grand Total |
| 5 | Count of Gender | | 20.31% 79.69% | 100.00% |

**Figure 1-24:** Gender breakdown of subscribers

You find that approximately 80 percent of subscribers are men, so ESPN may not want to advertise on *The View*!

## Describing the Income Distribution of Subscribers

In the `Income` worksheet (see Figure 1-25) you see a breakdown of the percentage of subscribers in each income group. This can be determined by performing the following steps:

1. In the `data` worksheet drag the Income field to the Row Labels zone and the Income field to the Values zone.

2. Change the Income field in the Values zone to Count of Income (if it isn't already there) and group incomes in $25,000 increments.
3. Finally, use Value Field Settings ➢ Show Value As ➢ % of Column Total to display the percentage of subscribers in each income bracket.

You see that a majority of subscribers are in the $54,000–$103,000 range. In addition, more than 85 percent of *ESPN: The Magazine* subscribers have income levels well above the national median household income, making them an attractive audience for ESPN's additional marketing efforts.

| | A | B |
|---|---|---|
| 1 | | |
| 2 | | |
| 3 | Row Labels ▾ | Count of Income |
| 4 | 29-53 | 13.18% |
| 5 | 54-78 | 35.06% |
| 6 | 79-103 | 29.20% |
| 7 | 104-128 | 14.55% |
| 8 | 129-153 | 5.66% |
| 9 | 154-178 | 1.56% |
| 10 | 179-203 | 0.59% |
| 11 | 229-253 | 0.10% |
| 12 | 254-278 | 0.10% |
| 13 | Grand Total | 100.00% |

**Figure 1-25:** Income distribution of subscribers

## Describing Subscriber Location

Next you will determine the breakdown of *ESPN: The Magazine* subscribers between suburbs, urban, and rural areas. This will help the analyst recommend the TV stations where ads should be placed.

1. Put your cursor inside the data from the data worksheet and drag the Location field to the Column Labels zone and Value zone.
2. Apply Value Field Settings and choose Show Value As ➢ % of Row Total to obtain the PivotTable, as shown in the Location worksheet and Figure 1-26. You find that 46 percent of subscribers live in the suburbs; 40 percent live in urban areas and 15 percent live in rural areas.

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | | | | | |
| 2 | | | | | |
| 3 | | Column Labels ▾ | | | |
| 4 | | rural | suburban | urban | Grand Total |
| 5 | Count of Location | 14.55% | 45.70% | 39.75% | 100.00% |

**Figure 1-26:** Breakdown of Subscriber Location

## Constructing a Crosstabs Analysis of Age and Income

Often marketers break down customer demographics simultaneously on two attributes. Such an analysis is called a *crosstabs* analysis. In the data worksheet you can perform a crosstabs analysis on Age and Income. To do so, perform the following steps:

1. Begin by dragging the Age field to the Column Labels zone and the Income field to the Row Labels and Value Labels zones.
2. Next group ages in 10-year increments and income in $25,000 increments.
3. Finally, use Value Field Settings to change the method of calculation to Show Value As ➤ % of Row Total. The Income and Age worksheet (shown in Figure 1-27) shows the resulting PivotTable, which indicates that 28.13 percent of subscribers in the $54,000 to $78,000 bracket are in the 28–37 age group.

Crosstabs analyses enable companies to identify specific combinations of customer demographics that can be used for a more precise allocation of their marketing investments, such as advertising and promotions expenditures. Crosstabs analyses are also useful to determine where firms *should not* make investments. For instance, there are hardly any subscribers to *ESPN: The Magazine* that are 78 or older, or with household incomes above $229,000, so placing ads on TV shows that are heavily watched by wealthy retiree is not recommended.

| Row Labels | 8-17 | 18-27 | 28-37 | 38-47 | 48-57 | 58-67 | 68-77 | 78-87 | 88-97 | 98-107 | 108-117 | Grand Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Count of Income | Column Labels | | | | | | | | | | | |
| 29-53 | 11.11% | 31.85% | 25.19% | 15.56% | 8.15% | 3.70% | 0.74% | 0.74% | 0.74% | 0.74% | 1.48% | 100.00% |
| 54-78 | 8.08% | 27.86% | 28.13% | 17.27% | 9.47% | 2.51% | 3.62% | 1.11% | 1.39% | 0.28% | 0.28% | 100.00% |
| 79-103 | 7.69% | 28.43% | 29.43% | 18.73% | 7.02% | 3.34% | 3.01% | 1.00% | 1.00% | 0.00% | 0.33% | 100.00% |
| 104-128 | 9.40% | 25.50% | 34.90% | 15.44% | 8.72% | 1.34% | 2.68% | 0.00% | 1.34% | 0.00% | 0.67% | 100.00% |
| 129-153 | 1.72% | 37.93% | 29.31% | 18.97% | 3.45% | 1.72% | 6.90% | 0.00% | 0.00% | 0.00% | 0.00% | 100.00% |
| 154-178 | 12.50% | 37.50% | 12.50% | 25.00% | 0.00% | 0.00% | 6.25% | 0.00% | 6.25% | 0.00% | 0.00% | 100.00% |
| 179-203 | 0.00% | 0.00% | 0.00% | 50.00% | 16.67% | 0.00% | 16.67% | 0.00% | 0.00% | 0.00% | 16.67% | 100.00% |
| 229-253 | 0.00% | 100.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100.00% |
| 254-278 | 0.00% | 0.00% | 100.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100.00% |
| Grand Total | 8.20% | 28.81% | 28.81% | 17.58% | 8.01% | 2.64% | 3.22% | 0.78% | 1.17% | 0.20% | 0.59% | 100.00% |

**Figure 1-27:** Crosstabs Analysis of subscribers

## Pulling Data from a PivotTable with the GETPIVOTDATA Function

Often a marketing analyst wants to pull data from a PivotTable to use as source information for a chart or other analyses. You can achieve this by using the GETPIVOTDATA function. To illustrate the use of the GETPIVOTDATA function, take a second look at

the True Colors hardware store data in the `products` worksheet from the `PARETO` `.xlsx` file.

1. With the cursor in any blank cell, type an = sign and point to the cell (B12) containing Adhesive 8 sales.

2. You will now see in the formerly blank cell the formula `=GETPIVOTDATA` `("Price",$A$3,"Product","Adhesive 8")`. Check your result against cell E10. This pulls the sales of Adhesive 8 ($42.00) from the PivotTable into cell E10.

This formula always picks out Adhesive 8 sales from the Price field in the PivotTable whose upper left corner is cell A3. Even if the set of products sold changes, this function still pulls Adhesive 8 sales.

In Excel 2010 or 2013 if you want to be able to click in a PivotTable and return a cell reference rather than a `GETPIVOTTABLE` function, simply choose File ➤ Options, and from the Formulas dialog box uncheck the `Use GetPivotData` functions for the PivotTable References option (see Figure 1-28).



| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | |
| 2 | | | | | | | | | |
| 3 | Row Labels | Revenue | | | | | | | |
| 4 | Adhesive 1 | $24.00 | | | | | | | |
| 5 | Adhesive 10 | $70.00 | | | | | | | |
| 6 | Adhesive 2 | $63.00 | | | | | | | |
| 7 | Adhesive 3 | $49.00 | | | | | | | |
| 8 | Adhesive 4 | $238.00 | | | | | | | |
| 9 | Adhesive 5 | $168.00 | | | | | | | |
| 10 | Adhesive 6 | $42.00 | | | 42 | | | | |
| 11 | Adhesive 7 | $70.00 | | | GETPIVOTDATA("Price",$A$3,"Product","Adhesive 8") | | | | |
| 12 | Adhesive 8 | $42.00 | | | | | | | |
| 13 | Adhesive 9 | $91.00 | | | | | | | |
| 14 | Safety 1 | $732.00 | | | | | | | |
| 15 | Safety 10 | $90.00 | | | | | | | |
| 16 | Safety 2 | $740.00 | | | | | | | |
| 17 | Safety 3 | $140.00 | | | | | | | |
| 18 | Safety 4 | $90.00 | | | | | | | |
| 19 | Safety 5 | $50.00 | | | | | | | |
| 20 | Safety 6 | $220.00 | | | | | | | |
| 21 | Safety 7 | $420.00 | | | | | | | |
| 22 | Safety 8 | $5,050.00 | | | | | | | |
| 23 | Safety 9 | $40.00 | | | | | | | |
| 24 | Tape 1 | $93.00 | | | | | | | |
| 25 | Tape 10 | $490.00 | | | | | | | |
| 26 | Tape 2 | $25.00 | | | | | | | |
| 27 | Tape 3 | $15.00 | | | | | | | |
| 28 | Tape 4 | $35.00 | | | | | | | |
| 29 | Tape 5 | $47.50 | | | | | | | |
| 30 | Tape 6 | $67.50 | | | | | | | |
| 31 | Tape 7 | $340.00 | | | | | | | |
| 32 | Tape 8 | $20.00 | | | | | | | |
| 33 | Tape 9 | $30.00 | | | | | | | |
| 34 | Grand Total | $9,592.00 | | | | | | | |

**Figure 1-28:** Example of GETPIVOTDATA function

This function is widely used in the corporate world and people who do not know it are at a severe disadvantage in making best use of PivotTables. Chapter 2 covers this topic in greater detail.

## Summary

In this chapter you learned the following:

- Sketch out in your mind how you want the PivotTable to look before you fill in the Field List.
- Use Value Field Settings to change the way the data is displayed or the type of calculation (Sum, Average, Count, etc.) used for a Value Field.
- A PivotChart can often clarify the meaning of a PivotTable.
- Double-click in a cell to drill down to the source data that was used in the cell's calculation.
- The GETPIVOTDATA function can be used to pull data from a PivotTable.

## Exercises

1. The Makeup2007.xlsx file (available for download on the companion website) gives sales data for a small makeup company. Each row lists the salesperson, product sold, location of the sale, units sold, and revenue generated. Use this file to perform the following exercises:

   a. Summarize the total revenue and units sold by each person of each product.

   b. Summarize the percentage of each person's sales that came from each location. Create a PivotChart to summarize this information.

   c. Summarize each girl's sales by location and use the Report Filter to change the calculations to include any subset of products.

2. The Station.xlsx file contains data for each family including the family size (large or small), income (high or low), and whether the family bought a station wagon. Use this file to perform the following exercises:

   a. Does it appear that family size or income is a more important determinant of station wagon purchases?

   b. Compute the percentage of station wagon purchasers that are high or low income.

    c.  Compute the fraction of station wagon purchasers that come from each of the following four categories: High Income Large Family, High Income Small Family, Low Income Large Family, and Low Income Small Family.

3.  The `cranberrydata.xlsx` file contains data for each quarter in the years 2006–2011 detailing the pounds of cranberries sold by a small grocery store. You also see the store's price and the price charged by the major competitor. Use this file to perform the following exercises:

    a.  Ignoring price, create a chart that displays the seasonality of cranberry sales.

    b.  Ignoring price, create a chart that shows whether there is an upward trend in sales.

    c.  Determine average sales per quarter, breaking it down based on whether your price was higher or lower than the competitor's price.

4.  The `tapedata.xlsx` file contains data for weeks during 2009–2011 for the unit sales of 3M tape, price charged, whether an ad campaign was run that week (1 = ad campaign), and whether the product was displayed on the end of the aisle (1 = end cap). Use this file to perform the following exercises:

    a.  Does there appear to be an upward trend in sales?

    b.  Analyze the nature of the monthly seasonality of tape sales.

    c.  Does an ad campaign appear to increase sales?

    d.  Does placing the tape in an end-cap display appear to increase sales?

5.  The files `EAST.xlsx` and `WEST.xlsx` contain information on product sales (products A–H) that you sell in January, February, and March. You want to use a PivotTable to summarize this sales data. The Field List discussed in this chapter does not enable you to create PivotTables on data from multiple ranges. If you hold down the ALT key and let go and hold down the D key and let go of the D key and hold down the P key, you can see the Classic PivotTable Wizard that enables you to select multiple ranges of data to key a PivotTable. Let Excel create a single page field for you and create a PivotTable to summarize total sales in the East and West during each month. Use the Filters so that the PivotTable shows only January and March sales of products A, C, and E.

# 2

# Using Excel Charts to Summarize Marketing Data

An important component of using analytics in business is a push towards "visualization." Marketing analysts often have to sift through mounds of data to draw important conclusions. Many times these conclusions are best represented in a chart or graph. As Confucius said, "a picture is worth a thousand words." This chapter focuses on developing Excel charting skills to enhance your ability to effectively summarize and present marketing data. This chapter covers the following skills:

- Using a combination chart and a secondary axis
- Adding a product picture to your column graphs
- Adding labels and data tables to your graphs
- Using the Table feature to ensure your graphs automatically update when new data is added
- Using PivotCharts to summarize a marketing survey
- Making chart labels dynamic
- Using Sparklines to summarize sales at different stores
- Using Custom Icon Sets to summarize trends in sales force performance
- Using a check box to control the data series that show in a graph
- Using the Table feature and `GETPIVOTDATA` function to automate creation of end-of-month sales reports

## Combination Charts

Companies often graph actual monthly sales and targeted monthly sales. If the marketing analyst plots both these series with straight lines it is difficult to differentiate between actual and target sales. For this reason analysts often summarize actual sales versus targeted sales via a combination chart in which the two series are displayed using different formats (such as a line and column graph.) This section explains how to create a similar combination chart.

All work for this chapter is located in the file `Chapter2charts.xlsx`. In the worksheet `Combinations` you see actual and target sales for the months January through July. To begin charting each month's actual and target sales, select the range F5:H12 and choose a line chart from the Insert tab. This yields a chart with both actual and target sales displayed as lines. This is the second chart shown in Figure 2-1; however, it is difficult to see the difference between the two lines here. You can avoid this by changing the format of one of the lines to a column graph. To do so, perform the following steps:

1. Select the Target sales series line in the line graph by moving the cursor to any point in this series.
2. Then right-click, and choose Change Series Chart Type...
3. Select the first Column option and you obtain the first chart shown in Figure 2-1.
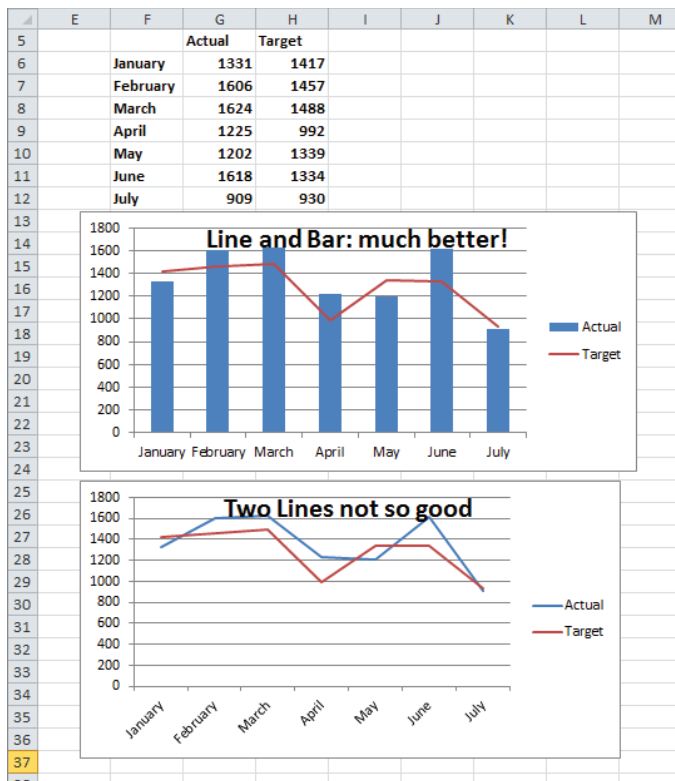


**Figure 2-1:** Combination chart

With the first chart, it is now easy to distinguish between the Actual and Target sales for each month. A chart like this with multiple types of graphs is called a *combination chart*.

You'll likely come across a situation in which you want to graph two series, but two series differ greatly in magnitude, so using the same vertical axis for each series results in one series being barely visible. In this case you need to include a secondary axis for one of the series. When you choose a secondary axis for a series Excel tries to scale the values of the axis to be consistent with the data values for that series. To illustrate the creation of a secondary axis, use the data in the `Secondary Axis` worksheet. This worksheet gives you monthly revenue and units sold for a company that sells expensive diamonds.

Because diamonds are expensive, naturally the monthly revenue is much larger than the units sold. This makes it difficult to see both revenue and units sold with only a single vertical axis. You can use a secondary axis here to more clearly summarize the monthly units sold and revenue earned. To do so, perform the following steps:

1. In the `Secondary Axis` worksheet select D7:F16, click the Insert tab, click the Line menu in the Charts section, and select the first Line chart option. You obtain the chart shown in Figure 2-2.
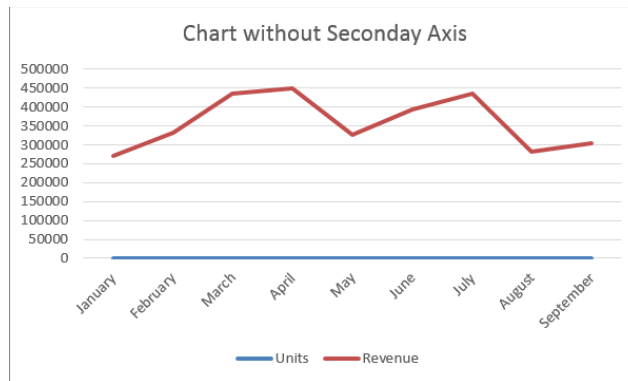


**Figure 2-2:** Line graph shows need for Secondary Axis

2. You cannot see the units' data though, because it is so small relative to the revenue. To remedy this problem, select the Revenue series line in the graph and right-click.

3. Choose Format Data Series, and select Secondary Axis.
4. Then select any point in the Revenue series, right-click, and change the chart type for the Revenue series to a column graph.

The resulting chart is shown in Figure 2-3, and you can now easily see how closely units and revenue move together. This occurs because Revenue=Average price*units sold, and if the average price during each month is constant, the revenue and units sold series will move in tandem. Your chart indeed indicates that average price is consistent for the charted data.
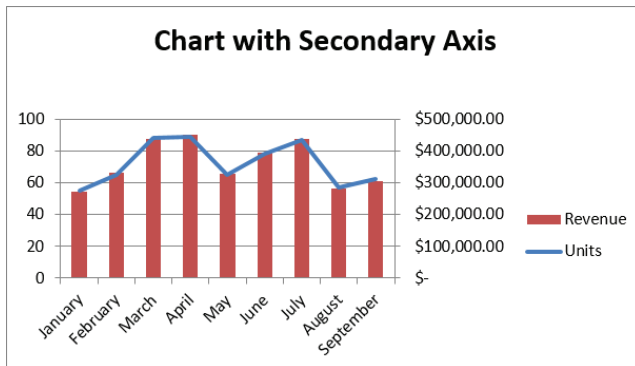


**Figure 2-3:** Chart with a secondary axis

## Adding Bling to a Column Graph with a Picture of Your Product

While column graphs are quite useful for analyzing data, they tend to grow dull day after day. Every once in a while, perhaps in a big presentation to a potential client, you might want to spice up a column graph here and there. For instance, in a column graph of your company's sales you could add a picture of your company's product to represent sales. Therefore if you sell Ferraris, you could use a .bmp image of a Ferrari in place of a bar to represent sales. To do this, perform the following steps:

1. Create a Column graph using the data in the `Picture` worksheet. This data shows monthly sales of cars at an L.A. Ferrari dealer.

2. After creating a Column graph, right-click on any column and choose Format Data Series... followed by Fill.
3. Click Picture or texture fill, as shown in Figure 2-4.
4. Next click File below the Insert from query and choose the `Ferrari.bmp` file (available for download from the companion site).
5. Choose Stack and Scale with **200** units per picture. This ensures each car represents sales of 200 cars.

The resulting chart is shown in Figure 2-5, as well as in the `Picture` worksheet.
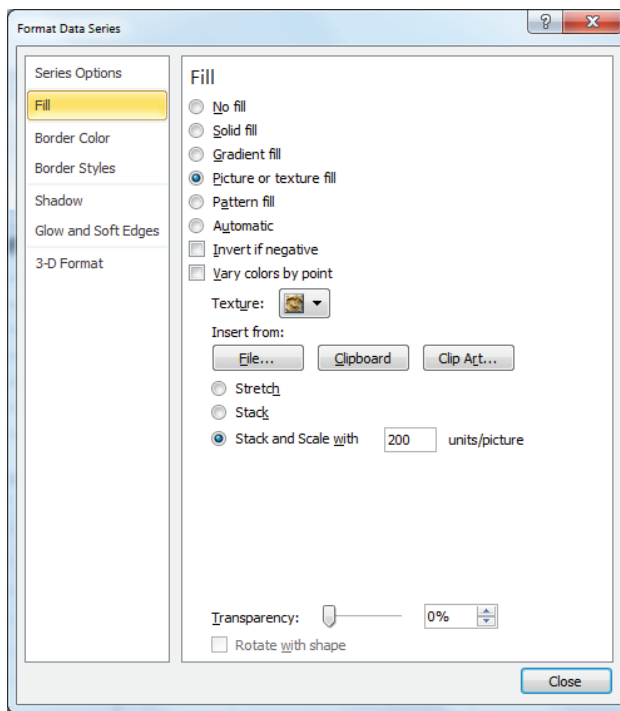


**Figure 2-4:** Dialog box for creating picture graph

**NOTE**    If you choose Stretch in the Fill tab of the Format Data Series dialog, you can ensure that Excel represents each month's sales by a single car whose size is proportional to that month's sales.
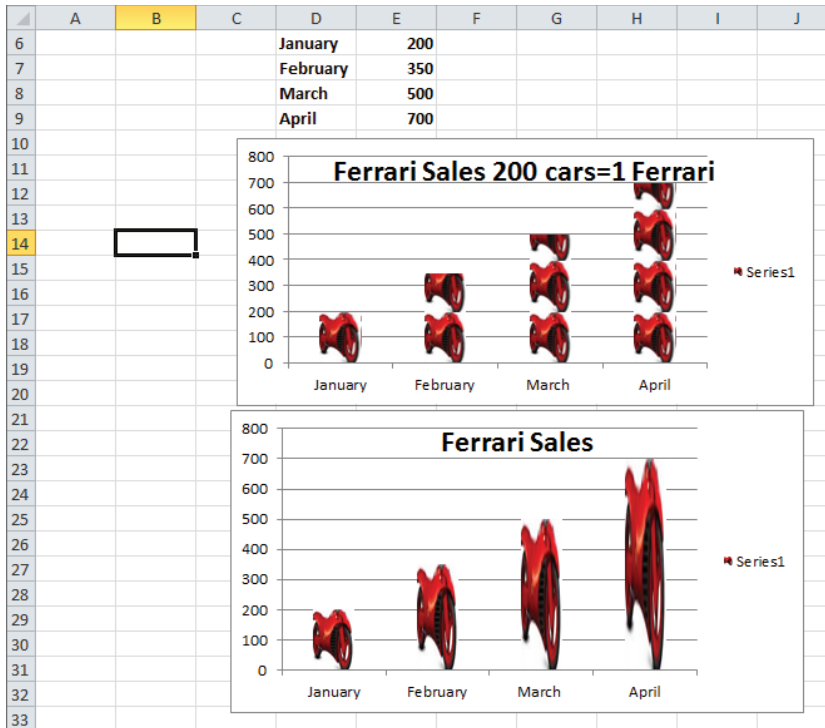
**Figure 2-5:** Ferrari picture graph

## Adding Labels or Tables to Your Charts

Often you want to add data labels or a table to a graph to indicate the numerical values being graphed. You can learn to do this by using example data that shows monthly sales of product lines at the True Color Department Store. Refer to the Labels and Tables worksheet for this section's examples.

To begin adding labels to a graph, perform the following steps:

1. Select the data range C5:D9 and choose the first Column chart option (Clustered Column) from the Charts section of the Insert tab.

2. Now click on any column in the graph and choose Layout from the Chart Tools section of the ribbon.

3. Choose the Data Labels option from the Layout section of Chart Tools and select More Data Label Options... Fill in the dialog boxes as shown in Figure 2-6.

4. Include the Value and Category Name in the label, and put them on a different line.

The resulting graph with data labels is shown in Figure 2-7.
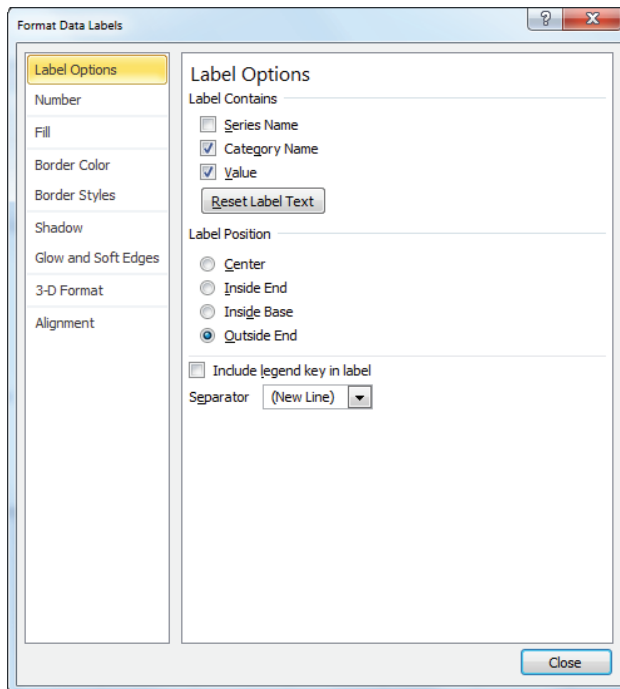


**Figure 2-6:** Dialog Box for creating chart with data labels

You can also add a Data Table to a column graph. To do so, again select any of the columns within the graph and perform the following steps:

1. From the Chart Tools tab, select Layout.

2. Choose Data Table.

3. Choose the Show Data Table option to create the table shown in the second chart in Figure 2-7.
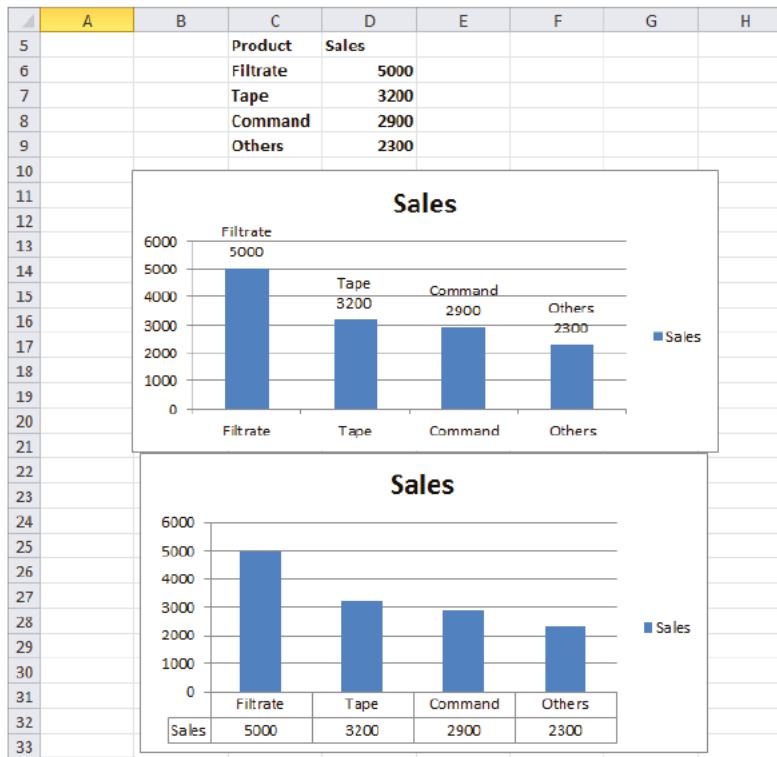
**Figure 2-7:** Chart with data labels or data tables

## Using a PivotChart to Summarize Market Research Surveys

To determine a new product's viability, market researchers often ask potential customers questions about a new product. PivotTables (discussed in Chapter 1, "Slicing and Dicing Marketing Data with PivotTables") are typically used to summarize this type of data. A PivotChart is a chart based on the contents of a PivotTable. As you will see, a PivotChart can be a highly effective tool to summarize any data gathered from this type of market research. The Survey PivotChart worksheet in the Chapter2chart.xlsx file uses a PivotChart to summarize a market research survey based on the data from the Survey Data worksheet (see Figure 2-8). The example data shows the answer to seven questions about a product (such as Likely to Buy) that was recorded on a 1–5 scale, with a higher score indicating a more favorable view of the product.

| | I | J | K |
|---|---|---|---|
| 5 | Response | Question | |
| 6 | 5 | Likely To Buy | |
| 7 | 3 | Easy to cook | |
| 8 | 3 | Attractive packaging | |
| 9 | 2 | Too expensive | |
| 10 | 5 | Easy to cook | |
| 11 | 1 | Easy to cook | |
| 12 | 3 | Too expensive | |
| 13 | 1 | Likely To Buy | |
| 14 | 2 | Likely To Buy | |
| 15 | 3 | Better tasting | |
| 16 | 2 | Too expensive | |
| 17 | 5 | Too expensive | |
| 18 | 3 | Nutritious meal | |
| 19 | 2 | Recommend to friend | |
| 20 | 2 | Nutritious meal | |
| 21 | 4 | Likely To Buy | |
| 22 | 2 | Likely To Buy | |
| 23 | 5 | Attractive packaging | |
| 24 | 4 | Likely To Buy | |
| 25 | 2 | Likely To Buy | |
| 26 | 5 | Easy to cook | |
| 27 | 5 | Attractive packaging | |
| 28 | 3 | Too expensive | |
| 29 | 4 | Recommend to friend | |
| 30 | 4 | Likely To Buy | |

**Figure 2-8:** Data for PivotChart example

To summarize this data, you can perform the following steps:

1. Select the data from the `Survey Data` worksheet and from the Tables section of the Insert tab choose PivotTable ➤ PivotTable and click OK.

2. Next drag the Question field heading to the Row Labels zone and the Response field heading to the Column Labels and Values zones.

3. Assume you want to chart the fraction of 1, 2, 3, 4, and 5 responses for each question. Therefore, use Value Field Settings to change the Summarize Values By tab field to Count and change the Show Values As tab field to % of Row Total. You then obtain the PivotTable shown first in Figure 2-9.

4. You can now create a Pivot Chart from this table. Select PivotChart from the Options tab and choose the first Column chart option. This process yields the PivotChart shown second in Figure 2-9.

5. This chart is a bit cluttered however, so click the drop-down arrow by the Question button, and choose to display only responses to Likely to Buy and Recommend to a Friend. The resulting uncluttered chart is shown in Figure 2-10.

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | |
| 2 | | | | | | | | |
| 3 | Count of Response | Column Labels ▼ | | | | | | |
| 4 | Row Labels ▼ | | 1 | 2 | 3 | 4 | 5 Grand Total | |
| 5 | Attractive packaging | | 8.33% | 6.55% | 16.07% | 37.50% | 31.55% | 100.00% |
| 6 | Better tasting | | 10.33% | 10.33% | 19.57% | 30.43% | 29.35% | 100.00% |
| 7 | Easy to cook | | 7.14% | 8.24% | 12.64% | 40.11% | 31.87% | 100.00% |
| 8 | Likely To Buy | | 9.78% | 9.78% | 14.13% | 37.50% | 28.80% | 100.00% |
| 9 | Nutritious meal | | 7.14% | 6.04% | 13.19% | 37.36% | 36.26% | 100.00% |
| 10 | Recommend to friend | | 9.47% | 8.95% | 11.58% | 33.16% | 36.84% | 100.00% |
| 11 | Too expensive | | 10.75% | 5.38% | 14.52% | 33.33% | 36.02% | 100.00% |
| 12 | Grand Total | | 9.01% | 7.92% | 14.50% | 35.58% | 32.99% | 100.00% |



**Figure 2-9:** PivotTable and cluttered PivotChart
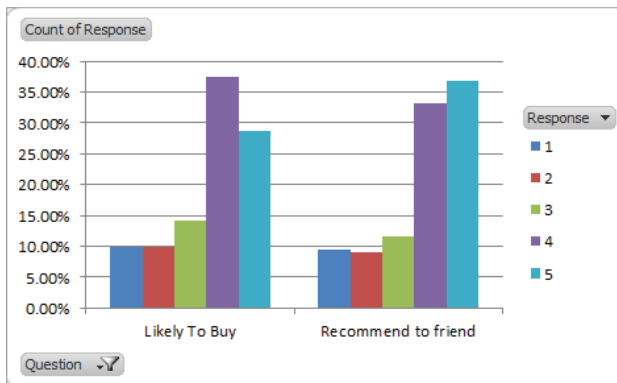


**Figure 2-10:** Uncluttered PivotChart

You can see that for both questions more than 60 percent of the potential customers gave the new product a 4 or 5, which is quite encouraging.

# Ensuring Charts Update Automatically When New Data is Added

Most marketing analysts download sales data at the end of each month and individually update a slew of graphs that include the new data. If you have Excel 2007 or later, an easy way to do this is to use the TABLE feature to cause graphs to automatically update to include new data. To illustrate the idea, consider the Housing starts worksheet, which contains U.S. monthly housing starts (in 000s) for the time period January 2000 through May 2011. From the Insert tab, you can create the X-Y scatter chart, as shown in Figure 2-11. Just click anywhere in the data and then choose the first scatter plot option.
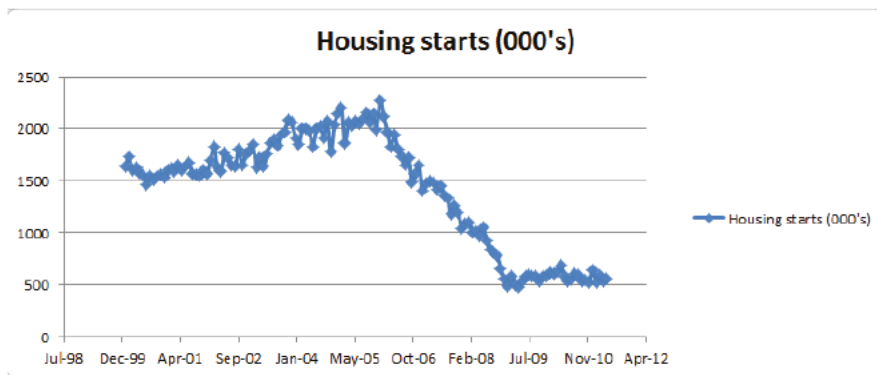


**Figure 2-11:** Housing chart through May 2011

If you add new data in columns D and E as it currently stands, the chart does not automatically incorporate the new data. To ensure that new data is automatically charted, you can instead make your data a TABLE before you create the chart. To do this, simply select the data (including headers) and press Ctrl+T. You can try this for yourself in the New Data worksheet by following these steps:

1. Select the data (D2:E141) and press Ctrl+T to make it a table.
2. Insert a scatter chart.
3. Add two new data points for June 2011 and July 2011 in Row 142 and Row 143.
4. The chart (as shown in Figure 2-12) automatically updates to include the new data. The difference between the two charts is apparent because June and July 2011 housing starts are much larger than April and May 2011 housing

starts. Imagine the time-savings if you have 15 or 20 charts keying off this data series!
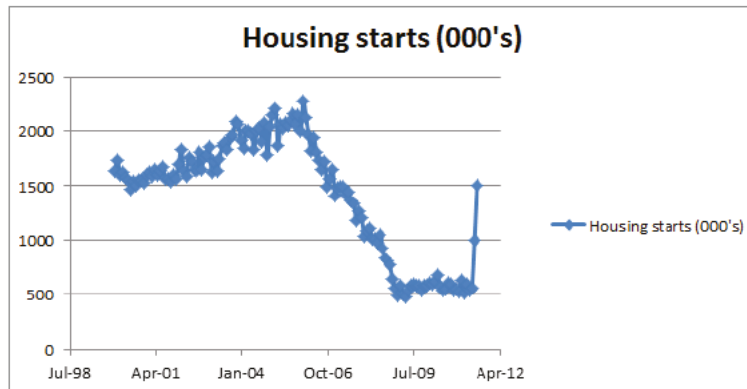


**Figure 2-12:** Housing chart through July 2011

You can use the Ctrl+T trick for PivotTables too. If you use data to create a PivotTable, you should make the data a TABLE before creating the PivotTable. This ensures that when you right-click and select Refresh (to update a PivotTable) the PivotTable automatically incorporates the new data.

# Making Chart Labels Dynamic

Data in your charts comes from data in your spreadsheet. Therefore changes to your spreadsheet will often alter your chart. If you do not set up the chart labels to also change, your chart will certainly confuse the user. This section shows how to make chart labels dynamically update to reflect changes in the chart.

Suppose you have been asked to project sales of a new product. If sales in a year are at least 2,000 units, the company will break even. There are two inputs in the Dynamic Labels worksheet: Year 1 sales (in cell D2) and annual growth in sales (in cell D3). The break-even target is given in cell D1. Use the Excel Create from Selection feature to name the data in D1:D3 with the names in the cell range C1:C3. To accomplish this, perform the following steps:

1. Select the cell range D1:D3.
2. Under the Formulas tab in the Defined Names section, select Create from Selection.

3. Choose the Left Column. Now, for example, when you use the name Year1sales in a formula, Excel knows you are referring to D2.

**NOTE**   A important reason for using range names is the fact that you can use the F3 key to Paste Range Names in formulas. This can save you a lot of time moving back and forth in your spreadsheet. For example, if your data is in Column A and you have named it you can use the F3 key to quickly and easily paste the data into a formula that is far away from Column A (for example, Column HZ.)

You can continue by graphing the annual sales that are based on initial sales and annual sales growth. As these assumptions change, the graph title and vertical axis legend need to reflect those changes. To accomplish this goal proceed as follows:

4. Enter the Year 1 sales in cell F5 with the formula `=Year1sales`. Copy the formula `=F5*(1+anngrowth)` from F6 to F7:F15 to generate sales in Years 2–11.

5. You want to find the year in which you break even, so copy the formula `=IF(F5>=target,"yes","no")` from G5 to G6:G15 to determine if you have broken even during the current year. For example, you can see that during Years 5–11 you broke even.

The key to creating dynamic labels that depend on your assumed initial sales and annual sales growth is to base your chart title and series legend key off formulas that update as your spreadsheet changes.

Now you need to determine the first year in which you broke even. This requires the use of the Excel `MATCH` function. The `MATCH` function has the following syntax: `MATCH(lookup_value,lookup_range,0)`. The `lookup_range` must be a row or column, and the `MATCH` function will return the position of the first occurrence of the `lookup_value` in the `lookup_range`. If the `lookup_value` does not occur in the `lookup_range`, the `MATCH` function returns an `#N/A` error. The last argument of 0 is necessary for the `MATCH` function to work properly.

In cell E19 the formula `=MATCH("yes",G5:G15,0)` returns the first year (in this case 5) in which you meet the breakeven target. In cell D22 several Excel functions are used to create the text string that will be the dynamic chart title. The formula entered in cell D22 is as follows:

```
=IFERROR("We will break even in year "&TEXT(E19,"0"),"Never break
even")
```

The functions used in this formula include the following:

- `IFERROR` evaluates the formula before the comma. If the formula does not evaluate to an error, `IFERROR` returns the formula's result. If the formula evaluates to an error, `IFERROR` returns what comes after the comma.
- The & or concatenate sign combines what is before the & sign with what comes after the & sign.
- The `TEXT` function converts the contents of a cell to text in the desired format. The two most commonly used formats are "0", which that indicates an integer format, and "0.0%", which indicates a percentage format.

If the `MATCH` formula in cell E19 does not return an error, the formula in cell D22 returns the text string We Will Break Even in Year followed by the first year you break even. If you never make it to break even, the formula returns the phrase Never Break Even. This is exactly the chart title you need.

You also need to create a title for the sales data that includes the annual sales growth rate. The series title is created in cell D23 with the following formula:

```
="Sales (growth rate="&TEXT(anngrowth,"0.0")&")".
```

**NOTE**   Note that the growth rate has been changed to a single decimal.

Now you are ready to create the chart with dynamic labels. To do so, perform the following steps:

1. Select the source data (cell range E4:F15 in the `Dynamic Labels` worksheet) from the chart.
2. Navigate to the Insert tab and choose an X-Y Scatter Chart (the second option).
3. Go to the Layout section of the Chart Tools Group and select Chart Title and Centered Overlay Chart.
4. Place the cursor in the formula bar and type an equal sign (=), point to the chart title in D22, and press Enter. You now have a dynamic chart title that depends on your sales assumptions.
5. To create the series title, right-click any plotted data point, and choose Select Data.
6. Click Edit and choose Series name.

7. Type an equals sign, point to the title in D23, and press Enter. You now have a dynamic series label. Figure 2-13 shows the resulting chart.
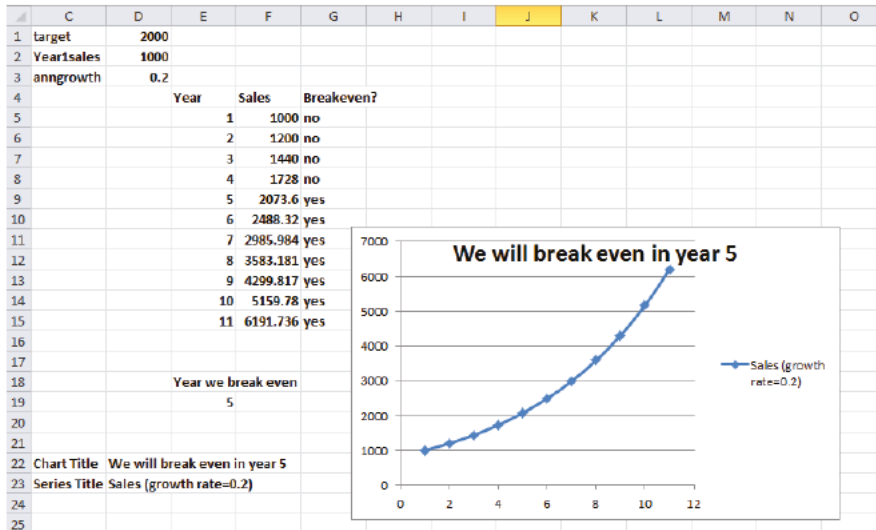


**Figure 2-13:** Chart with dynamic labels

## Summarizing Monthly Sales-Force Rankings

If you manage a sales force, you need to determine if a salesperson's performance is improving or declining. Creating customized icon sets provides an easy way to track a salesperson's performance over time.

The data in the Sales Tracker worksheet list sales of salespeople during each month (see Figure 2-14). You can track each month with icons (up arrow, down arrow, or flat arrow) to determine whether a salesperson's ranking has improved, worsened, or stayed the same. You could use Excel's Conditional Formatting icon sets (see Chapter 24 of Winston's *Data Analysis and Business Modeling with Excel 2010* for a description of the Conditional Formatting icon sets), but then you would need to insert a set of icons separately for each column. A more efficient (although not as pretty) way to create icons is to use Wingdings 3 font and choose an "h" for an up arrow, an "i" for a down arrow, and a "g" for a flat arrow (see Figure 2-15).

| | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | p | h | up | | | | | | | | | | | | |
| 2 | q | i | down | | | | | | | | | | | | |
| 3 | u | g | flat | | | | | | | | | | | | |
| 4 | | | | | | | rank | rank | rank | rank | rank | trend | trend | trend | trend |
| 5 | | January | February | March | April | May | January | February | March | April | May | Feb | March | April | May |
| 6 | Lebron | 85 | 66 | 81 | 61 | 56 | 10 | 14 | 13 | 15 | 15 ↓ | ↑ | ↓ | → |
| 7 | Wade | 82 | 63 | 74 | 78 | 75 | 12 | 15 | 15 | 12 | 14 ↓ | → | ↑ | ↓ |
| 8 | Dirk | 45 | 100 | 115 | 127 | 150 | 15 | 4 | 3 | 2 | 1 ↑ | ↑ | ↑ | ↑ |
| 9 | Manning | 75 | 88 | 89 | 76 | 83 | 13 | 10 | 11 | 13 | 12 ↑ | ↓ | ↓ | ↑ |
| 10 | Brady | 96 | 90 | 98 | 76 | 93 | 9 | 9 | 9 | 13 | 10 → | ↓ | ↓ | ↑ |
| 11 | Halliday | 75 | 73 | 79 | 91 | 95 | 13 | 13 | 14 | 10 | 9 → | ↓ | ↑ | ↑ |
| 12 | Britney | 98 | 91 | 109 | 99 | 84 | 8 | 8 | 4 | 8 | 11 → | ↑ | ↓ | ↓ |
| 13 | Lindsay | 83 | 84 | 97 | 81 | 98 | 11 | 12 | 10 | 11 | 8 ↓ | ↑ | ↓ | ↑ |
| 14 | Paris | 106 | 98 | 84 | 93 | 82 | 5 | 5 | 12 | 9 | 13 → | ↓ | ↑ | ↓ |
| 15 | JLO | 104 | 88 | 109 | 101 | 115 | 6 | 10 | 4 | 6 | 3 ↓ | ↑ | ↓ | ↑ |
| 16 | Emma | 115 | 94 | 105 | 101 | 107 | 3 | 7 | 6 | 6 | 6 ↓ | ↑ | → | → |
| 17 | Melo | 118 | 98 | 128 | 126 | 108 | 2 | 5 | 1 | 3 | 4 ↓ | ↑ | ↓ | ↓ |
| 18 | KD | 100 | 114 | 104 | 116 | 131 | 7 | 2 | 7 | 5 | 2 ↑ | ↓ | ↑ | ↑ |
| 19 | Vick | 112 | 122 | 102 | 124 | 107 | 4 | 1 | 8 | 4 | 6 ↑ | ↓ | ↑ | ↓ |
| 20 | Rodgers | 127 | 114 | 116 | 139 | 108 | 1 | 2 | 2 | 1 | 4 ↓ | → | ↑ | ↓ |
| 21 | | | | | | | | | | | | | | | |

Figure 2-14: Monthly sales data

| J | K | L |
|---|---|---|
| Letter | Wingdings 3 | |
| a | ⇨ | |
| b | ⇦ | |
| c | ⇨ | |
| d | ⇦ | |
| e | ⇨ | |
| f | ← | |
| g | → | |
| h | ↑ | |
| i | ↓ | |
| j | ↖ | |
| k | ↗ | |
| l | ↙ | |
| m | ↘ | |
| n | ↔ | |
| o | ↕ | |
| p | ▲ | |
| q | ▼ | |
| r | △ | |
| s | ▽ | |
| t | ◀ | |
| u | ▶ | |
| v | ◁ | |
| w | ▷ | |
| x | ◣ | |
| y | ◢ | |
| z | ◤ | |

Figure 2-15: Icon sets created with Wingdings 3 font

To begin creating up, down, or flat arrows that reflect monthly changes in each salesperson's performance, perform the following steps:

1. Create each salesperson's rank during each month by copying the formula
   =RANK(E6,E$6:E$20,0) from J6 to J6:N20.

The last argument of 0 in the RANK function indicates that the largest number in E6:E20 will receive a rank of 1, the second largest number a rank of 2, and so on. If the last argument of the RANK function is 1, the smallest number receives a RANK of 1.

2. Next create the h, i, and g that correspond to the arrows by copying the formula =IF(K6<J6,"h",IF(K6>J6,"i","g")) from O6 to O6:R20.
3. Finally change the font in O6:R20 to Wingdings3.

You can now follow the progress of each salesperson (relative to her peers). For example, you can now see that Dirk Nowitzki improved his ranking each month.

## Using Check Boxes to Control Data in a Chart

Often the marketing analyst wants to plot many series (such as sales of each product) in a single chart. This can result in a highly cluttered chart. Often it is convenient to use *check boxes* to control the series that are charted. The Checkboxes worksheet illustrates this idea. The data here shows weekly sales of chocolates, DVDs, magazines, soda, and hot dogs at the Quickie Mart convenience store. You can easily summarize this data in a single chart with five lines, as shown in Figure 2-16
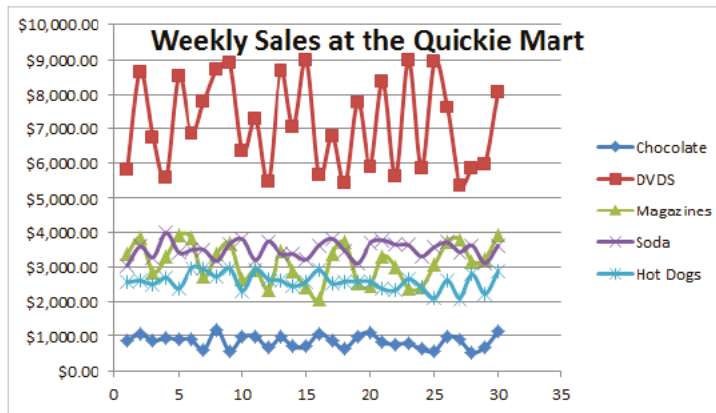


**Figure 2-16:** Sales at Quickie Mart

Now this chart is a good start, however, you might notice that the presence of five series appears cluttered, and therefore you would probably be better off if you

did not show the sales of every product. Check boxes make it easy to control which series show up in a chart. To create a check box, perform the following steps:

1. Ensure you see the Developer tab on the ribbon. If you do not see the Developer tab in Excel 2010 or 2013, select File ➢ Options, and from Customize Ribbon select the Developer tab.

2. To place a check box in a worksheet, select the Developer tab and chose Insert.

3. From the Form Controls (Not ActiveX) select the check box that is third from the left in the top row.

4. After releasing the left mouse in the worksheet, you see a drawing tool. Hold down the left mouse to size the check box. If you want to resize the check box again later, simply hold down the control key and click the check box. A check box simply puts a TRUE or FALSE in a cell of your choosing. It serves as a Toggle switch that can be used to turn Excel features (such as a Function argument or Conditional Formatting) on or off.

5. Next, select the cell controlled by the check box by moving the cursor to the check box until you see the Pointer (a hand with a pointing finger).

6. Right-click, select the Format Control dialog box, and select cell F3 (by point-ing) as the cell link. Click OK.

7. Now when this check box is selected, F3 shows TRUE, and when the check box is not selected, cell F3 shows FALSE. Label this check box with the name of the series it controls, Chocolate in this case.

In a similar fashion four more check boxes were created that control G3, H3, I3, and J3. These check boxes control the DVDs, magazines, sodas, and hot dogs series, respectively. Figure 2-17 shows these check boxes.

After you have created the necessary check boxes, the trick is to not chart the original data but chart data controlled by the check boxes. The key idea is that Excel will ignore a cell containing an #N/A error when charting. To get started, copy the formula = `IF(F$3,F6,NA())` from O6 to O6:S35 to ensure that a series will only be charted when its check box is checked. For instance, check the Chocolate check box and a TRUE appears in F3, so O6 just picks up the actual week 1 chocolate sales in F6. If you uncheck the Chocolate check box, you get a FALSE in F3, O6 returns #N/A, and nothing shows up in the chart. To chart the data in question, select the modified data in the range O5:S35 and choose from the Insert tab the second X Y Scatter option. If you check only the Chocolate and Magazines check boxes, you

can obtain the chart shown in Figure 2-18, which shows only sales of Chocolate and Magazines.

| | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | |
| 2 | | | | | | | | | |
| 3 | | | | | TRUE | FALSE | TRUE | FALSE | FALSE |
| 4 | | | | | Original Data | | | | |
| 5 | | | | Week | Chocolate | DVDS | Magazines | Soda | Hot Dogs |
| 6 | | | | 1 | $857.00 | $5,820.00 | $3,374.00 | $3,036.00 | $2,564.00 |
| 7 | | | | 2 | $1,050.00 | $8,656.00 | $3,821.00 | $3,589.00 | $2,615.00 |
| 8 | ☑ Chocolate | | | 3 | $863.00 | $6,749.00 | $2,857.00 | $3,279.00 | $2,512.00 |
| 9 | | | | 4 | $933.00 | $5,580.00 | $3,284.00 | $3,981.00 | $2,705.00 |
| 10 | | | | 5 | $905.00 | $8,534.00 | $3,923.00 | $3,411.00 | $2,386.00 |
| 11 | ☐ DVDS | | | 6 | $901.00 | $6,837.00 | $3,837.00 | $3,484.00 | $2,977.00 |
| 12 | | | | 7 | $597.00 | $7,794.00 | $2,713.00 | $3,498.00 | $2,946.00 |
| 13 | | | | 8 | $1,185.00 | $8,708.00 | $3,375.00 | $3,169.00 | $2,729.00 |
| 14 | | | | 9 | $565.00 | $8,892.00 | $3,674.00 | $3,660.00 | $2,977.00 |
| 15 | | | | 10 | $964.00 | $6,338.00 | $2,640.00 | $3,797.00 | $2,322.00 |
| 16 | ☑ Magazines | | | 11 | $982.00 | $7,258.00 | $2,926.00 | $3,201.00 | $2,939.00 |
| 17 | | | | 12 | $658.00 | $5,468.00 | $2,360.00 | $3,738.00 | $2,648.00 |
| 18 | | | | 13 | $961.00 | $8,690.00 | $3,466.00 | $3,374.00 | $2,619.00 |
| 19 | ☐ Soda | | | 14 | $691.00 | $7,054.00 | $2,888.00 | $3,371.00 | $2,449.00 |
| 20 | | | | 15 | $695.00 | $8,988.00 | $2,433.00 | $3,211.00 | $2,589.00 |
| 21 | | | | 16 | $1,065.00 | $5,658.00 | $2,085.00 | $3,622.00 | $2,919.00 |
| 22 | | | | 17 | $851.00 | $6,781.00 | $3,392.00 | $3,812.00 | $2,533.00 |
| 23 | ☐ Hot dogs | | | 18 | $636.00 | $5,423.00 | $3,714.00 | $3,458.00 | $2,590.00 |
| 24 | | | | 19 | $989.00 | $7,763.00 | $2,534.00 | $3,101.00 | $2,596.00 |

**Figure 2-17:** Check boxes for controlling which products show in chart
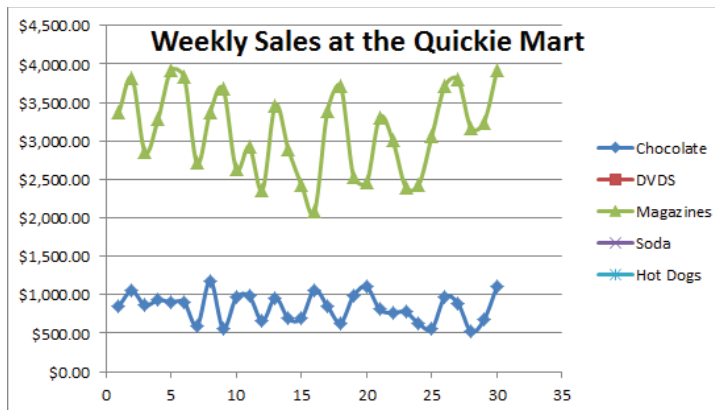


**Figure 2-18:** Charting only selected data

# Using Sparklines to Summarize Multiple Data Series

Suppose you are charting daily sales of French Fries at each of the over 12,000 US McDonald's. Showing the sales for each restaurant in a single chart would result in a useless, cluttered graph. But think of the possibilities if you could summarize daily sales for each restaurant in a single cell! Fortunately, Excel 2010 and later enables you to create *sparklines*. A sparkline is a graph that summarizes a row or column of data in a single cell. This section shows you how to use Excel 2010 to easily create sparklines.

The Sparklines worksheet contains data that can be used to illustrate the concept of Sparklines. The data gives the number of engagement rings sold each day of the week in each city for a national jewelry store chain (see Figure 2-19).

| | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | |
| 2 | | Different scale for each sparkline | | | | | | | |
| 3 | | Monday | Tuesday | Wednesday | Thursday | Friday | Saturday | Sunday | |
| 4 | Chicago | 2520 | 3680 | 3852 | 3816 | 4800 | 5668 | 3600 | |
| 5 | New York | 6804 | 9630 | 8262 | 7290 | 10044 | 11700 | 9630 | |
| 6 | Los Angeles | 6489 | 8640 | 8910 | 8181 | 11340 | 12753 | 8550 | |
| 7 | San Francisco | 2184 | 3000 | 2673 | 2646 | 3528 | 4134 | 2970 | |
| 8 | Dallas | 3325 | 4650 | 4050 | 4050 | 6540 | 6630 | 5200 | |
| 9 | Houston | 4494 | 6120 | 5724 | 5778 | 6840 | 8580 | 5700 | |
| 10 | Atlanta | 3052 | 4240 | 3384 | 3924 | 4512 | 4680 | 4400 | |
| 11 | Seattle | 1890 | 3300 | 2700 | 2619 | 3960 | 3939 | 3030 | |
| 12 | Miami | 1316 | 1820 | 1800 | 1728 | 2328 | 2678 | 2120 | |
| 13 | Boston | 3948 | 6180 | 5562 | 5670 | 7056 | 7488 | 5940 | |
| 14 | Nashville | 1498 | 2020 | 1818 | 1962 | 2352 | 2600 | 1880 | |
| 15 | | | | | | | | | |
| 16 | | Same Scale for each sparkline | | | | | | | |
| 17 | | Monday | Tuesday | Wednesday | Thursday | Friday | Saturday | Sunday | |
| 18 | Chicago | 2520 | 3680 | 3852 | 3816 | 4800 | 5668 | 3600 | |
| 19 | New York | 6804 | 9630 | 8262 | 7290 | 10044 | 11700 | 9630 | |
| 20 | Los Angeles | 6489 | 8640 | 8910 | 8181 | 11340 | 12753 | 8550 | |
| 21 | San Francisco | 2184 | 3000 | 2673 | 2646 | 3528 | 4134 | 2970 | |
| 22 | Dallas | 3325 | 4650 | 4050 | 4050 | 6540 | 6630 | 5200 | |
| 23 | Houston | 4494 | 6120 | 5724 | 5778 | 6840 | 8580 | 5700 | |
| 24 | Atlanta | 3052 | 4240 | 3384 | 3924 | 4512 | 4680 | 4400 | |
| 25 | Seattle | 1890 | 3300 | 2700 | 2619 | 3960 | 3939 | 3030 | |
| 26 | Miami | 1316 | 1820 | 1800 | 1728 | 2328 | 2678 | 2120 | |
| 27 | Boston | 3948 | 6180 | 5562 | 5670 | 7056 | 7488 | 5940 | |
| 28 | Nashville | 1498 | 2020 | 1818 | 1962 | 2352 | 2600 | 1880 | |

**Figure 2-19:** Sparklines Example

You can summarize the daily sales by graphing the daily counts for each city in a single cell. To do so, perform the following steps:

1. First, select where you want your Sparklines to go (the `Sparklines` worksheet uses K4:K14: you can use L4:L14) and then from the Insert tab, select Line from the Sparklines section.

2. Fill in the dialog box shown in Figure 2-20 with the data range on which the Sparklines are based (D4:J14).
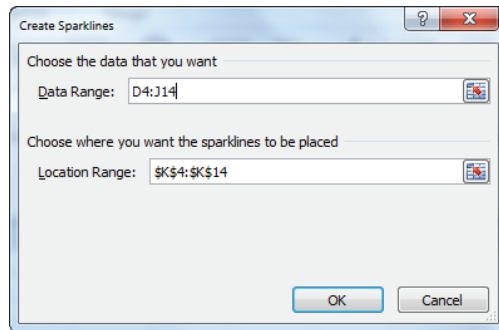


**Figure 2-20:** Sparkline dialog box

You now see a line graph (refer to Figure 2-19) that summarizes the daily sales in each city. The Sparklines make it clear that Saturday is the busiest day of the week.

If you click in any cell containing a Sparkline, the Sparkline Tools Design tab appears. Here, after selecting the Design tab, you can make many changes to your Sparklines. For example, Figure 2-21 shows selections for the high and low points to be marked and Figure 2-22 shows the Sparklines resulting from these selections.
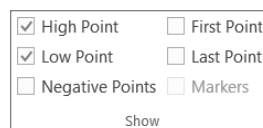


**Figure 2-21:** Selecting High and Low Point Markers

| | N | O | P | Q | R | S | T | U | V |
|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | |
| 2 | | Different scale for each sparkline | | | | | | | |
| 3 | | Monday | Tuesday | Wednesd: | Thursday | Friday | Saturday | Sunday | |
| 4 | Chicago | 2520 | 3680 | 3852 | 3816 | 4800 | 5668 | 3600 | |
| 5 | New York | 6804 | 9630 | 8262 | 7290 | 10044 | 11700 | 9630 | |
| 6 | Los Angel | 6489 | 8640 | 8910 | 8181 | 11340 | 12753 | 8550 | |
| 7 | San Franci | 2184 | 3000 | 2673 | 2646 | 3528 | 4134 | 2970 | |
| 8 | Dallas | 3325 | 4650 | 4050 | 4050 | 6540 | 6630 | 5200 | |
| 9 | Houston | 4494 | 6120 | 5724 | 5778 | 6840 | 8580 | 5700 | |
| 10 | Atlanta | 3052 | 4240 | 3384 | 3924 | 4512 | 4680 | 4400 | |
| 11 | Seattle | 1890 | 3300 | 2700 | 2619 | 3960 | 3939 | 3030 | |
| 12 | Miami | 1316 | 1820 | 1800 | 1728 | 2328 | 2678 | 2120 | |
| 13 | Boston | 3948 | 6180 | 5562 | 5670 | 7056 | 7488 | 5940 | |
| 14 | Nashville | 1498 | 2020 | 1818 | 1962 | 2352 | 2600 | 1880 | |

**Figure 2-22:** Sparklines with High and Low Markers

These Sparklines make it clear that Saturday was the busiest day for each branch and Monday was the slowest day.

The Design tab enables you to make the following changes to your Sparklines:

- Alter the type of Sparkline (Line, Column, or Win-Loss). Column and Win-Loss Sparklines are discussed later in the chapter.
- Use the Edit Data choice to change the data used to create the Sparkline. You can also change the default setting so that hidden data is included in your Sparkline.
- Select any combination of the high point, low point, negative points, first point, or last point to be marked.
- Change the style or color associated with the Sparklines and markers.
- Use the Axis menu to change the way the axes are set for each Sparkline. For example, you may make the x-axis or y-axis scale the same for each Sparkline. This is the scaling used in the cell range K18:K28. Note this choice shows that Nashville, for example, sells fewer rings than the other cities. The default is for the scale for each axis to be based on the data for the individual Sparkline. This is the scaling used in the Sparklines in the cell range K4:K14 of Figure 2-22. The Custom Value choice enables you to pick an upper and lower limit for each axis.
- When data points occur at irregularly spaced dates, you can select Data Axis Type from the Axis menu so that the graphed points are separated by an amount of space proportional to the differences in dates.

**NOTE**   By clicking any of the Sparklines, you could change them to Column Sparklines simply by selecting Column from the Sparklines Design tab.

Excel also can create Win-Loss Sparklines. A Win-Loss Sparkline treats any positive number as an *up block* and any negative number as a *down block*. Any 0s are graphed as a gap. A Win-Loss Sparkline provides a perfect way to summarize performance against sales targets. In the range D32:J42 you can see a list of the daily targets for each city. By copying the formula = IF(D18>D32−1−1) from D45 to D45:J55, you can create a 1 when a target is met and *a* −1 when a target is not met. To create the Win-Loss Sparklines, select the range where the Sparklines should be placed (cell range K45:K55) and from the Insert menu, select Win-Loss Sparklines. Then choose the data range D45:J55. Figure 2-23 shows your Win-Loss Sparklines.

| | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|
| 30 | | **Target** | | | | | | | |
| 31 | | Monday | Tuesday | Wednesday | Thursday | Friday | Saturday | Sunday | |
| 32 | Chicago | 3193 | 3592 | 3592 | 3991 | 3991 | 5188 | 3991 | |
| 33 | New York | 7241 | 8146 | 8146 | 9051 | 9051 | 11767 | 9051 | |
| 34 | Los Angeles | 7413 | 8340 | 8340 | 9266 | 9266 | 12046 | 9266 | |
| 35 | San Francisco | 2415 | 2717 | 2717 | 3019 | 3019 | 3925 | 3019 | |
| 36 | Dallas | 3937 | 4429 | 4429 | 4921 | 4921 | 6397 | 4921 | |
| 37 | Houston | 4941 | 5559 | 5559 | 6177 | 6177 | 8030 | 6177 | |
| 38 | Atlanta | 3222 | 3625 | 3625 | 4027 | 4027 | 5236 | 4027 | |
| 39 | Seattle | 2450 | 2756 | 2756 | 3063 | 3063 | 3981 | 3063 | |
| 40 | Miami | 1576 | 1773 | 1773 | 1970 | 1970 | 2561 | 1970 | |
| 41 | Boston | 4782 | 5380 | 5380 | 5978 | 5978 | 7771 | 5978 | |
| 42 | Nashville | 1615 | 1817 | 1817 | 2019 | 2019 | 2624 | 2019 | |
| 43 | | Monday | Tuesday | Wednesday | Thursday | Friday | Saturday | Sunday | |
| 44 | | **Met Goal?** | | | | | | | |
| 45 | Chicago | -1 | 1 | 1 | -1 | 1 | 1 | -1 | |
| 46 | New York | -1 | 1 | 1 | -1 | 1 | -1 | 1 | |
| 47 | Los Angeles | -1 | 1 | 1 | -1 | 1 | 1 | -1 | |
| 48 | San Francisco | -1 | 1 | -1 | -1 | 1 | 1 | -1 | |
| 49 | Dallas | -1 | 1 | -1 | -1 | 1 | 1 | 1 | |
| 50 | Houston | -1 | 1 | 1 | -1 | 1 | 1 | -1 | |
| 51 | Atlanta | -1 | 1 | -1 | -1 | 1 | -1 | 1 | |
| 52 | Seattle | -1 | 1 | -1 | -1 | 1 | -1 | -1 | |
| 53 | Miami | -1 | 1 | 1 | -1 | 1 | 1 | 1 | |
| 54 | Boston | -1 | 1 | 1 | -1 | 1 | -1 | -1 | |
| 55 | Nashville | -1 | 1 | 1 | -1 | 1 | -1 | -1 | |

**Figure 2-23:** Win-Loss Sparklines

**NOTE**   If you want your Sparklines to automatically update to include new data, you should make the data a table.

# Using GETPIVOTDATA to Create the End-of-Week Sales Report

Many marketing analysts download weekly sales data and want to summarize it in charts that update as you download the new data. The Excel Table feature and GETPIVOTDATA function can greatly ease this process. To illustrate this process, download the source data in columns D through G in the End of Month dashboard worksheet. Each row (see Figure 2-24) gives the week of sales, the sales category, the store in which sales were made, and the revenue generated. The goal is to create a line graph to summarize sales of selected product categories (for each store) that automatically update when new data is downloaded. The approach is as follows:

| | D | E | F | G |
|---|---|---|---|---|
| 4 | WEEK | Catego | Store | Revenue |
| 5 | 3 | Abrasives | Lowe's | 2043 |
| 6 | 12 | Safety | Menards | 2343 |
| 7 | 3 | Tape | Home Depo | 1414 |
| 8 | 12 | Command | Target | 1820 |
| 9 | 9 | Tape | Home Depo | 943 |
| 10 | 7 | Tape | Target | 1219 |
| 11 | 7 | Command | Menards | 1156 |
| 12 | 11 | Abrasives | Lowe's | 2127 |
| 13 | 12 | Safety | Menards | 1315 |
| 14 | 3 | Tape | Target | 1580 |
| 15 | 10 | Abrasives | Home Depo | 1598 |
| 16 | 4 | Command | Lowe's | 1000 |
| 17 | 7 | Tape | Menards | 1087 |
| 18 | 7 | Abrasives | Menards | 1728 |
| 19 | 1 | Abrasives | Target | 1911 |
| 20 | 7 | Abrasives | Menards | 1563 |
| 21 | 2 | Tape | Target | 2482 |
| 22 | 7 | Safety | Lowe's | 1534 |
| 23 | 12 | Safety | Menards | 1471 |
| 24 | 2 | Abrasives | Lowe's | 990 |
| 25 | 11 | Tape | Lowe's | 1580 |
| 26 | 1 | Safety | Target | 2389 |

**Figure 2-24:** Weekly sales data

1. Make the source data a table; in this example you can do so by selecting Table from the Insert tab after selecting the range D4:G243.
2. Create a PivotTable based on the source data by dragging Week to the Row Labels zone, Store and Category to the Column Labels zone, and Revenue to the Values zone. This PivotTable summarizes weekly sales in each store for each product category. Figure 2-25 shows a portion of this PivotTable.

| | J | K | L | M | N | O | P | Q | R | S | T | U |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 11 | Column Labels ▼ | | | | | | | | | | | |
| 12 | Lowe's | | | | Lowe's Total | Menards | | | | Menards Total | Home Depot | |
| 13 | Abrasives | Safety | Tape | Command | | Abrasives | Safety | Tape | Command | | Abrasives | Safety |
| 14 | | 1065 | | 1247 | 2312 | 3772 | 526 | 3466 | | 7764 | 3272 | 1984 |
| 15 | 2349 | 4510 | | | 6859 | 6220 | 3296 | 2159 | | 11675 | 1921 | 2167 |
| 16 | 2797 | 1233 | 2046 | 6335 | 12411 | 3472 | | 959 | 2369 | 6800 | | 3352 |
| 17 | 1501 | 3822 | 1180 | 1000 | 7503 | 6018 | 2212 | | | 8230 | | |
| 18 | 1940 | | 833 | 2945 | 5718 | 708 | | 501 | 597 | 1806 | 1758 | 3755 |
| 19 | | | 1918 | 6494 | 8412 | 4145 | 6205 | 4653 | | 15003 | | 2253 |
| 20 | 3049 | 3228 | | | 6277 | 3291 | 2303 | 6472 | 1156 | 13222 | 3580 | 4218 |
| 21 | 4208 | 2890 | 1873 | | 8971 | 910 | | 1494 | 3776 | 6180 | 2445 | 5496 |
| 22 | | | 2052 | 2825 | 4877 | 6265 | | 1888 | | 8153 | | |
| 23 | 6117 | | 1875 | | 7992 | | | 2138 | | 2138 | 4807 | 2325 |
| 24 | 2127 | 1971 | 4461 | 2394 | 10953 | 3596 | 1169 | | | 4765 | | 810 |
| 25 | | 1542 | 4786 | 850 | 7178 | 1124 | 7565 | 7264 | 2431 | 18384 | 2471 | |
| 26 | | | | 1196 | 1196 | | | | 3400 | 3400 | | |
| 27 | 1455 | | | | 1455 | | | | | | | |
| 28 | 25543 | 20261 | 21024 | 25286 | 92114 | 39521 | 23276 | 30994 | 13729 | 107520 | 20254 | 26360 |

**Figure 2-25:** PivotTable for Weekly sales report

3. Create a drop-down box in cell AG8 from which a store can be selected. To accomplish this, navigate to the Data tab and select Data Validation from the Data Tools group.

4. Choose Settings and fill in the Data Validation dialog box, as shown in Figure 2-26. This ensures that when the cursor is placed in cell AG8 you can see a drop-down box that lists the store names (pulled from cell range F5:F8)
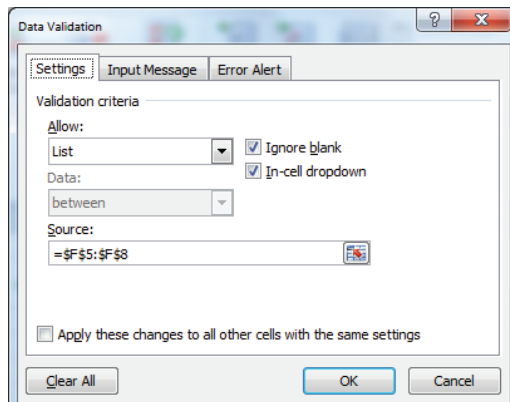


**Figure 2-26:** Creating Data Validation drop-down box

5. Enter week numbers in the range AG11:AG24 and product categories in AG10:AK10. Then use the GETPIVOTDATA function to pull the needed data from the PivotTable. Before entering the key formula, simply click anywhere in the PivotTable to get a GETPIVOTDATA function. Then copy that function and

paste it in the following formula to modify the arguments for Week, Category, and Store, as shown here. Enter this important formula in cell AH11:

```
=IF(AH$9=FALSE,NA(),IFERROR(GETPIVOTDATA("Revenue",$I$11,"WEEK"
,$AG11,"Category",AH$10,"Store",$AG$8)," ")).
```

6. Copy this formula to the range AH11:AK24. This formula uses GETPIVOTDATA to pull revenue for the store listed in AG8 and the category listed in row 10 if row 9 has a TRUE. The use of IFERROR ensures a blank is entered if the category were not sold in the selected store during the given week.

7. Make the source data for the chart, AG10:AK24, a table, so new data is automatically included in the graph. The Table feature can also ensure that when you enter a new week the GETPIVOTDATA formulas automatically copy down and pull needed information from the updated PivotTable.

8. Use check boxes to control the appearance of TRUE and FALSE in AH9:AK9. After making the range AG10:AK24 a table, the chart is simply an X-Y scatter graph with source data AG10:AK24. Figure 2-27 shows the result.
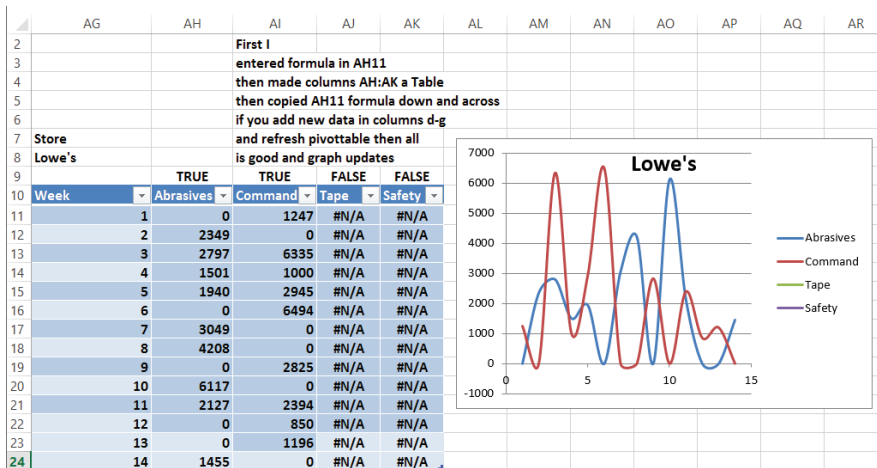


**Figure 2-27:** Sales summary report

9. Now add new data for Week 15 and refresh the PivotTable. If you add Week 15 in cell AG25, the graph automatically updates. Because you made the source data for the chart a table, when you press Enter Excel "knows" you want the formulas copied down.

## Summary

In this chapter you learned the following:

- Right-clicking a series in a chart enables you to change the series' chart type and create a Combination chart.
- Right-clicking a chart series and choosing Format Data Series enables you to create a secondary axis for a chart.
- Right-clicking a chart series in a Column graph and selecting Fill followed by Picture or Texture Fill enables you to replace a bland column with a picture from a file or Clip Art.
- On Chart Tools from the Layout tab, you can easily insert Data Labels and a Data Table in the chart.
- If you make the source data for a chart a table (from the Insert tab choose Table) before creating a chart, the chart automatically updates to include new data.
- PivotCharts are a great way to summarize a lengthy market research survey. Filtering on the questions enables you customize the results shown in your chart.
- If you base your chart title and series labels on cell formulas, they dynamically update as you change the inputs or assumptions in your spreadsheet.
- Clever use of IF formulas and the Wingdings 3 font can enable you to create a visually appealing summary of trends over time in sales data.
- Use check boxes to control the series that appear in a chart.
- Combining the Table feature, PivotTables, `GETPIVOTDATA`, check boxes, and Data Validation drop-down boxes makes it easy to create charts with customized views that automatically update to include new data.

## Exercises

Exercises 1–6 use data in the file `Chapter2exercisesdata.xlsx`.

1. The `Weather` worksheet includes monthly average temperature and rainfall in Bloomington, Indiana. Create a combination chart involving a column and line graph with a secondary axis to chart the temperature and rainfall data.
2. The `Weather` worksheet includes monthly average temperature and rainfall in Bloomington, Indiana. Create a combination chart involving a column and area graph with a secondary axis to chart the temperature and rainfall data.

3. The `Pictures and Labels` worksheet includes monthly tomato sales on Farmer Smith's farm. Summarize this data with pictures of tomatoes, data labels, and a data table.

4. The `Survey` worksheet contains results evaluating a training seminar for salespeople. Use a PivotChart to summarize the evaluation data.

5. The data in the `checkboxes` worksheet contains monthly sales during 2010 and 2011. Use check boxes to set up a chart in which the user can choose which series are charted.

6. The `Income` worksheet contains annual data on median income in each state for the years 1984–2010. Use Sparklines to summarize this data.

7. Jack Welch's GE performance evaluation system requires management to classify the top 20 percent of all workers, middle 70 percent of all workers, and bottom 10 percent of all workers. Use Icon sets to classify the salespeople in the `Sales Tracker` worksheet of file `Chapter2data.xlsx` according to the 20-70-10 rule.

**NOTE**   You need the `PERCENTILE` function. For example, the function `PERCENTILE(A1:A50,.9)` would return the 90th percentile of the data in the cell range A1:A50.

Exercises 8 and 9 deal with the data in the file `LaPetitbakery.xlsx` that was discussed in Chapter 1.

8. Use the Excel Table feature to set up a chart of daily cake sales that updates automatically when new data is included.

9. Set up a chart that can be customized to show total monthly sales for any subset of La Petit Bakery's products. Of course, you want the chart to update automatically when new data appears in the worksheet.

10. The marketing product life cycle postulates that sales of a new product will increase for a while and then decrease. Specify the following five inputs:

   - Year 1 sales
   - Years of growth
   - Years of decline
   - Average annual growth rate during growth period
   - Average annual decline during decline period

Set up a Data Validation drop-down box that allows years of growth and decline to vary between 3 and 10. Then determine sales during Years 1–20. Suppose 10,000 units need to be sold in a year to break even. Chart your annual sales, and set up a dynamic chart title that shows the year (if any) in which you break even. Your series label should include the values of the five input parameters.

# 3

# Using Excel Functions to Summarize Marketing Data

I n Chapter 1, "Slicing and Dicing Marketing Data with PivotTables," and Chapter 2, "Using Excel Charts to Summarize Marketing Data," you learned how to summarize marketing data with PivotTables and charts. Excel also has a rich library of powerful functions that you can use to summarize marketing data. In this chapter you will learn how these Excel functions enable marketers to gain insights for their data that can aid them to make informed and better marketing decisions.

This chapter covers the following:

- Using the Excel Table feature to create dynamic histograms that summarize marketing data and update automatically as new data is added.
- Using Excel statistical functions such as AVERAGE, STDEV, RANK, PERCENTILE, and PERCENTRANK to summarize marketing data.
- Using the powerful "counting and summing functions" (COUNTIF, COUNTIFS, SUMIF, SUMIFS, AVERAGEIF, and AVERAGEIFS) to summarize marketing data.
- Writing array formulas to perform complicated statistical calculations on any subset of your data.

## Summarizing Data with a Histogram

A *histogram* is a commonly used tool to summarize data. A histogram essentially tells you how many data points fall in various ranges of values. Several Excel functions, including *array functions* such as the TRANSPOSE and the FREQUENCY function, can be used to create a histogram that automatically updates when new data is entered into the spreadsheet.

When using an array function, you need to observe the following rules:

1.  Select the range of cells that will be populated by the array function. Often an array function populates more than one cell, so this is important.
2.  Type in the function just like an ordinary formula.
3.  To complete the entry of the formula, do not just press Enter; instead press Control+Shift+Enter. This is called *array entering* a formula.

Because the FREQUENCY function is a rather difficult array function, let's begin with a discussion of a simpler array function, the TRANSPOSE function.

## Using the TRANSPOSE Function

The TRANSPOSE function is a great place to start when illustrating the use of array functions. To begin, take a look at the Histogram worksheet in the Chapter3bakery. xlsx file. In Figure 3-1, cells N27:Q27 list some of my best students. This list would be easier to digest though if the names were listed in a column. To do this, copy N27:Q27 and move the cursor to R28, then right-click and select Paste Special... Select the Transpose checkbox to paste the names in a column.

| | N | O | P | Q | R | S |
|---|---|---|---|---|---|---|
| | | | | | PASTE | TRANSPOSE |
| 27 | Scarlett | Lindsay | Britney | Paris | SPECIAL | FUNCTION |
| 28 | | | | | Scarlett | Scarlett |
| 29 | | | | | Lindsay | Lindsay |
| 30 | | | | | Britney | Britney |
| 31 | | | | | Paris | Paris |

**Figure 3-1:** Use of the Transpose function

Unfortunately, if you change the source data (for example, Scarlett to Blake) the transposed data in column R does not reflect the change. If you want the transposed data to reflect changes in the source data, you need to use the TRANSPOSE function. To do so perform the following steps:

1.  Select S28:S31.
2.  Enter in cell S28 the formula =TRANSPOSE(N27:Q27) and press Control+Shift+Enter. You will know an array formula has been entered when you see a { in the Formula bar.
3.  Change Scarlett in N27 to Blake, and you can see that Scarlett changes to Blake in S28 but not in R28.

## Using the FREQUENCY Function

The `FREQUENCY` function can be used to count how many values in a data set fit into various ranges. For example, given a list of heights of NBA basketball players the `FREQUENCY` function could be used to determine how many players are at least 7' tall, how many players are between 6' 10" and 7' tall, and so on. The syntax of the `FREQUENCY` function is `FREQUENCY(array,bin range)`. In this section you will utilize the `FREQUENCY` function to create a histogram that charts the number of days in which cake sales fall into different numerical ranges. Moreover, your chart will automatically update to include new data.

On the `Histogram` worksheet you are given daily sales of cakes at La Petit Bakery. You can summarize this data with a histogram. The completed histogram is provided for you in the worksheet, but if you want to follow along with the steps simply copy the data in D7:E1102 to the same cell range in a blank worksheet. Begin by dividing a range of data into 5–15 equal size ranges (called a *bin range*) to key a histogram. To do so, perform the following steps:

1. Select your data (the range D7:E1102) and make this range a table by selecting Insert ➢ Table. If a cell range is already a table then under Insert the table option will be grayed out.
2. In cell E5, point to all the data and compute the minimum daily cake sales with the formula `=MIN(Table1[Cakes])`.
3. Similarly, determine the maximum daily cake sales in E6. The range of daily cake sales is between 32 and 165. Select your bin ranges to be <= 30, 31–50, 51–70, …, 151–170…, > 170 cakes.
4. List the boundaries for these bin ranges in the cell range I9:I17. Figure 3-2 shows this process in action.

You can now use the `FREQUENCY` function to count the number of days in which cake sales fall in each bin range. The syntax of the `FREQUENCY` function is `FREQUENCY(array,bin range)`. When an array is entered, this formula can count how many values in the array range fall into the bins defined by the bin range. In this example the function will count the number of days in which cake sales fall in each bin range. This is exemplified in Figure 3-2. To begin using the `FREQUENCY` function, perform the following steps:

1. Select the range J9:J17 in the `Histogram` worksheet and type the formula `=FREQUENCY(E8:E1102,I9:I17)`.

2.  Array enter this formula by pressing Control+Shift+Enter. This formula computes how many numbers in the table are ≤30, between 31 and 50,…, between 151 and 170, and more than 170.

| | D | E | F | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|
| 5 | min | 32 | | | | | | skewness | 0.427162 | |
| 6 | max | 165 | | | | | | | | |
| 7 | Date | Cakes | | | | | | | | |
| 8 | 1/1/2013 | 79 | | | | | | | | |
| 9 | 1/2/2013 | 91 | | | | 30 | 0 | | | |
| 10 | 1/3/2013 | 47 | | | | 50 | 27 | | | |
| 11 | 1/4/2013 | 89 | | | | 70 | 207 | | | |
| 12 | 1/5/2013 | 112 | | | | 90 | 454 | | | |
| 13 | 1/6/2013 | 89 | | | | 110 | 270 | | | |
| 14 | 1/7/2013 | 70 | | | | 130 | 109 | | | |
| 15 | 1/8/2013 | 70 | | | | 150 | 25 | | | |
| 16 | 1/9/2013 | 59 | | | | 170 | 3 | | | |
| 17 | 1/10/2013 | 71 | | | | >170 | 0 | | | |
| 18 | 1/11/2013 | 74 | | | | | | | | |
| 19 | 1/12/2013 | 119 | | | | | | | | |
| 20 | 1/13/2013 | 90 | | | | | | | | |
| 21 | 1/14/2013 | 96 | | | | | | | | |
| 22 | 1/15/2013 | 62 | | | | | | | | |
| 23 | 1/16/2013 | 48 | | | | | | | | |
| 24 | 1/17/2013 | 58 | | | | | | | | |
| 25 | 1/18/2013 | 74 | | | | | | | | |
| 26 | 1/19/2013 | 94 | | | | | | | | |
| 27 | 1/20/2013 | 97 | | | | | | | | |
| 28 | 1/21/2013 | 79 | | | | | | | | |
| 29 | 1/22/2013 | 67 | | | | | | | | |

**Figure 3-2:** Dynamic histogram for cake sales

For example, you can find that there were three days in which between 151 and 170 cakes were sold. The counts update automatically if new days of sales are added.

3.  Select the range I9:J17, and from the Insert tab, select the first Column Graph option.
4.  Right-click any column, select Format Data Series…, and change Gap Width to 0. This enables you to obtain the histogram, as shown in Figure 3-2.
5.  If you enter more data in Column E (say 10 numbers >170) you can see that the histogram automatically updates to include the new data. This would not have happened if you had failed to make the data a Table.

## Skewness and Histogram Shapes

There are a variety of histograms that are seen when examining marketing data, but the three most commonly seen histogram shapes are listed here (see Figure 3-3 and file `Skewexamples.xlsx`):

- *Symmetric histogram:* A histogram is symmetric and looks approximately the same to the left of the peak and the right of the peak. In Figure 3-3 the IQ scores yield an asymmetric histogram.

- *Positively skewed:* (skewed right) A histogram is positively skewed (or skewed right) if it has a single peak and the values of the data set extend much further to the right of the peak than to the left of the peak. In Figure 3-3 the histogram of family income is positively skewed.

- *Negatively skewed:* (skewed left) A histogram is negatively skewed if it has a single peak and the values of the data set extend much further to the left of the peak than to the right of the peak. In Figure 3-3, the histogram of Days from Conception to Birth is negatively skewed.
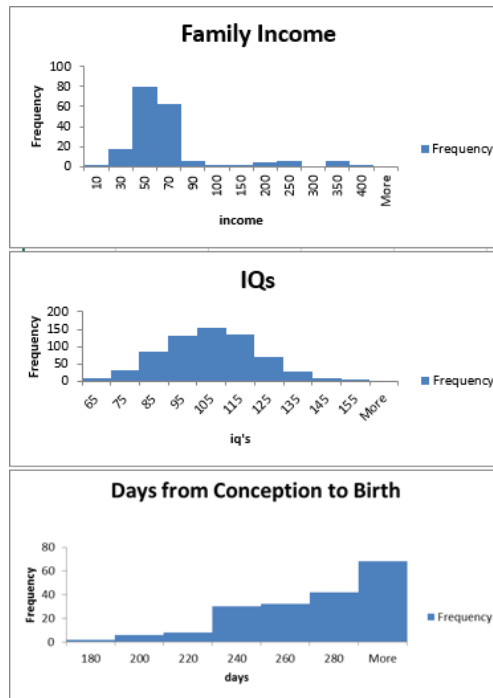


**Figure 3-3:** Symmetric, positively skewed, and negatively skewed histograms

The Excel SKEW function provides a measure of the skewness of a data set. A skewness value > +1 indicates mostly positive skewness, < –1 mostly negative skewness, and a skewness value between –1 and +1 indicates a mostly symmetric histogram. This function can be used to obtain a quick characterization of whether a histogram is symmetric, positively skewed, or negatively skewed. To see this function in action, enter the formula =SKEW(Table1[Cakes]) into cell L5 in the Histogram worksheet to yield a skewness of .43, implying that cake cells follow a symmetric distribution. Because you made the source data an Excel table, the skewness calculation in L5 would automatically update if new data is added.

In the next section you will see that the degree of skewness (or lack thereof) in a data set enables the marketing analyst to determine how to best describe the typical value in a data set.

## Using Statistical Functions to Summarize Marketing Data

In her daily work the marketing analyst will often encounter large datasets. It is often difficult to make sense of these data sets because of their vastness, so it is helpful to summarize the data on two dimensions:

- A typical value of the data: For example, what single number best summarizes the typical amount a customer spends on a trip to the supermarket.
- Spread or variation about the typical value: For example, there is usually much more spread about the typical amount spent on a visit to a grocery superstore than a visit to a small convenience store.

Excel contains many statistical functions that can summarize a data set's typical value and the spread of the data about its mean. This section discusses these statistical functions in detail.

### Using Excel Functions to Compute the Typical Value for a Data Set

The file Chapter3bakery.xlsx contains the La Petit Bakery sales data that was discussed in Chapter 1. For each product, suppose you want to summarize the typical

number of that product sold in a day. Three measures are often used to summarize the *typical value* for a data set.

- The mean, or *average*, is simply the sum of the numbers in the data set divided by the number of values in the dataset. The average is computed with the Excel AVERAGE function. For example, the average of the numbers 1, 3, and 5 is $\frac{1+3+5}{3}$ = 3. You can often use $\bar{x}$ to denote the mean or average of a data set.
- The *median* is approximately the 50th percentile of the data set. Roughly one-half the data is below the median and one-half the data is above the median. More precisely, suppose there are *n* values that when ordered from smallest to largest are $x_1, x_2, \ldots, x_n$. If *n* is odd the median is $x_{.5(n+1)}$, and if *n* is even, the median is $(x_{.5n} + x_{.5n+1})/2$ . For example, for the data set 1, 3, 5, the median is 3, and for the data set 1, 2, 3, 4 the median is 2.5. The MEDIAN function can be used to compute the median for a data set.
- The *mode* of a data set is simply the most frequently occurring value in the data set. If no value appears more than once, the data set has no mode. The MODE function can be used to compute the mode of a data set. A data set can have more than one mode. If you have Excel 2010 or later, then the array function MODE.mult (see *Excel 2010 Data Analysis and Business Modeling* by Wayne Winston [Microsoft Press, 2011]) can be used to compute all modes. For example, the data set 1, 3, and 5 has no mode, whereas the data set 1, 2, 2, 3 has a mode of 2.

The Descriptive stats worksheet uses Excel functions to compute three measures of typical value for daily sales of cakes, pies, cookies, smoothies, and coffee. Complete the following steps and see Figure 3-4 for details.

1. To compute the average for daily cake sales in cell H5, enter the formula =AVERAGE(H12:H1106).
2. To compute the median for daily cake sales in cell H6, enter the formula =MEDIAN(H12:H1106).
3. To, compute the mode of daily cake sales in cell H7, enter the formula =MODE(H12:H1106).
4. Copy these formulas to the range I5:L7 to compute these measures for La Petit Bakery's other products.

You find, for example, that on average 86.28 cakes were sold each day, around one-half the time fewer than 85 cakes were sold, and the most frequently occurring number of cakes sold in a day is 90.

| | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|
| 1 | 5th largest cake day | 148 | | | | | |
| 2 | 3rd smallest cake day | 38 | | | | | |
| 3 | rank of 136 cakes | 0.979 | | | | | |
| 4 | | | | | | | |
| 5 | mean | 86.28037 | 53.54795 | 540.4511 | 218.4666667 | 391.6265 | |
| 6 | median | 85 | 52 | 533 | 206 | 388 | |
| 7 | mode | 90 | 46 | 397 | 166 | 375 | |
| 8 | standard deviation | 20.27667 | 13.37189 | 132.3544 | 74.35093017 | 102.9305 | |
| 9 | 95th percentile | 122 | 77.2 | 764.8 | 352.2 | 568.4 | |
| 10 | Skewness | 0.427162 | 0.474088 | 0.440287 | 0.785999912 | 0.360745 | |
| 11 | Date | Cakes | Pies | Cookies | Smoothies | Coffee | promotion |
| 12 | 1/1/2013 | 79 | 46 | 518 | 60 | 233 | none |
| 13 | 1/2/2013 | 91 | 50 | 539 | 161 | 427 | none |
| 14 | 1/3/2013 | 47 | 60 | 222 | 166 | 347 | none |
| 15 | 1/4/2013 | 89 | 64 | 734 | 153 | 358 | none |
| 16 | 1/5/2013 | 112 | 73 | 764 | 240 | 392 | none |
| 17 | 1/6/2013 | 89 | 57 | 922 | 259 | 510 | none |
| 18 | 1/7/2013 | 70 | 50 | 476 | 120 | 334 | none |
| 19 | 1/8/2013 | 70 | 48 | 496 | 222 | 316 | none |
| 20 | 1/9/2013 | 59 | 37 | 587 | 181 | 156 | none |
| 21 | 1/10/2013 | 71 | 36 | 488 | 178 | 298 | none |
| 22 | 1/11/2013 | 74 | 50 | 645 | 100 | 490 | none |
| 23 | 1/12/2013 | 119 | 71 | 438 | 162 | 416 | none |
| 24 | 1/13/2013 | 90 | 51 | 568 | 137 | 434 | none |
| 25 | 1/14/2013 | 96 | 48 | 585 | 194 | 573 | promotion |

**Figure 3-4:** Descriptive statistics

## Which Measure of Typical Value Is Best?

Although there is no single best answer to this question, there are a few rules of thumb that are summarized here. The mode might be relevant if a hat store wanted to maximize sales and was going to stock only one size of hats. In most cases, however, the analyst who wants to summarize a data set by a single number must choose either the mean or the median. In the presence of substantial positive or negative skewness, extreme values tend to distort the mean, and the median is a better choice as a summary of a typical data value. In other situations using the median throws out important information, so the mean should be used to summarize a data set. To summarize:

- If a skew statistic is greater than 1 or less than −1, substantial skewness exists and you should use median as a measure of typical value.
- If skew statistic is between −1 and 1, use mean as a measure of typical value.

For the cake example, data skewness is between −1 and 1, so you would use the mean of 86.28 cakes as a summary value for daily cake sales.

For an example of how skewness can distort the mean, consider the starting salaries of North Carolina students who graduated in 1984. Geography majors had by far the highest average starting salary. At virtually every other school, business majors had the highest average starting salary. You may guess that Michael Jordan was a UNC geography major, and his multimillion dollar salary greatly distorted the average salary for geography majors. The median salary at UNC in 1984 was, as expected, highest for business majors.

## Using the VAR and STDEV Functions to Summarize Variation

Knowledge of the typical value characterizing a data set is not enough to completely describe a data set. For example, suppose every week Customer 1 spends $40 on groceries and during a week Customer 2 is equally likely to spend $10, $20, $30, or $100 on groceries. Each customer spends an average of $40 per week on groceries but we are much less certain about the amount of money Customer 2 will spend on groceries during a week. For this reason, the description of a data set is not complete unless you can measure the spread of a data set about its mean. This section discusses how the variance and standard deviation can be used to measure the spread of a data set about its mean.

Given a data set $x_1, x_2, \ldots, x_n$ the *sample variance* ($s^2$) of a data set is approximately the average squared deviation of the mean and may be computed with the formula $s^2 = \frac{1}{n-1} \sum_{i=1}^{i=n} (x_i - \bar{x})^2$. Here $\bar{x}$ is the average of the data values. For example, for the data set 1, 3, 5 the average is 3, so the sample variance is as follows:

$$\frac{1}{2}\{(1 - 3)^2 + (3 - 3)^2 + (5 - 3)^2\} = 4$$

Note that if all data points are identical, then the sample variance is 0. You should square deviations from the mean to ensure that positive and negative deviations from the mean will not cancel each other out.

Because variance is not in the same units as dollars, the analyst will usually take the square root of the sample variance to obtain the *sample standard deviation(s)*. Because the data set 1, 3, and 5, has a sample variance of 4, s=2.

The Excel `VAR` function computes sample variance, and the Excel `STDEV` function computes sample standard deviation. As shown in Figure 3-4 (see `Descriptive stats` worksheet), copying the formula `=STDEV(H12:H1106)` from H8 to K8:L8 computes the standard deviation of each product's daily sales. For example, the standard deviation of daily cake sales was 20.28 cakes.

## The Rule of Thumb for Summarizing Data Sets

If a data set does not exhibit substantial skewness, then the statisticians' interesting and important rule of thumb lets you easily characterize a data set. This rule states the following:

- Approximately 68 percent of all data points are within 1 sample standard deviation of the mean.
- Approximately 95 percent of all data points are within 2 standard deviations (2*s*) of the mean.
- Approximately 99.7 percent of the data points are within 3 standard deviations of the mean.

Any data point that is more than two sample standard deviations (2*s*) from the mean is unusual and is called an *outlier*. Later in the chapter (see the "Verifying the Rule of Thumb" section) you can find that for daily cake sales, the rule of thumb implies that 95 percent of all daily cake sales should be between 45.3 and 126.83. Thus cake sales less than 45 or greater than 127 would be an outlier. You can also find that 95.62 percent of all daily cake sales are within 2*s* of the mean. This is in close agreement with the rule of thumb.

## The Percentile.exc and Percentrank.exc Functions

A common reorder policy in a supply chain is to produce enough to have a low percent chance, say, a 5 percent, of running out of a product. For cakes, this would imply that you should produce *x* cakes, where there is a 5 percent chance demand for cakes that exceed *x* or a 95 percent chance demand for cakes that is less than or equal to *x*. The value *x* is defined to be the 95th percentile of cake demand. The Excel 2010 and later function `PERCENTILE.EXC(range, k)` (here *k* is between 0 and 1) returns the 100\**k*th percentile for data in a range. In earlier versions of Excel, you should use the `PERCENTILE` function. By copying the formula `=PERCENTILE.EXC(H12:H1106,0.95)` from H9 to I9:L9 of the `Descriptive stats` worksheet you can obtain the 95th percentile of daily sales for each product. For example, there is a 5 percent chance that more than 122 cakes will be sold in a day.

Sometimes you want to know how unusual an observation is. For example, on December 27, 2015, 136 cakes were sold. This is more than two standard deviations above average daily cake sales, so this is an outlier or unusual observation. The `PERCENTRANK.EXC` function in Excel 2010 or later (or `PERCENTRANK` in earlier versions of Excel) gives the ranking of an observation relative to all values in a data

set. The syntax of `PERCENTRANK.EXC` is `PERCENTRANK(range, x,[significance])`. This returns the percentile rank (as a decimal) of *x* in the given range. Significance is an optional argument that gives the number of decimal points returned by the function. Entering in cell H3 the formula `=PERCENTRANK.EXC(H12:H1106,H1102)`, you find that you sold more than 136 cakes on only 2.1 percent of all days.

## The LARGE and SMALL Functions

Often you want to find, say, the fifth largest or third smallest value in an array. The function `LARGE(range,k)` returns the *k*th largest number in a range, whereas the function `SMALL(range,k)` returns the *k*th smallest number in a range. As shown in Figure 3-4, entering in cell H1 the formula `=LARGE(H12:H1106,5)` tells you the fifth largest daily sales of cakes was 148. Similarly, entering in cell H2 the formula `=SMALL(H12:H1106,3)` tells you the third smallest daily demand for cakes was 38.

These same powerful statistical functions can be used by marketing managers to gain important insights such as who are the 5 percent most profitable customers, who are the three most costly customers, or even what percentage of customers are unprofitable.

## Using the COUNTIF and SUMIF Functions

In addition to mathematical functions such as `SUM` and `AVERAGE`, the `COUNTIF` and `SUMIF` functions might be (or should be!) two of the most used functions by marketing analysts. These functions provide powerful tools that enable the analyst to select a subset of data in a spreadsheet and perform a calculation (`count`, `sum`, or `average`) on any column in the spreadsheet. The wide variety of computations performed in this section should convince the marketer that these functions are a necessary part of your toolkit.

The first few examples involve the `COUNTIF` and `SUMIF` functions. The `COUNTIF` function has the syntax `COUNTIF(range, criteria)`. Then the `COUNTIF` function counts how many cells in the range meet the criteria. The `SUMIF` function has the syntax `SUMIF (range, criteria,` and `sum_range)`. Then the `SUMIF` function adds up the entries in the sumrange column for every row in which the cell in the range column meets the desired criteria.

The following sections provide some examples of these functions in action. The examples deal with the La Petit Bakery data, and the work is in the `Sumif Countif` worksheet. The work is also shown in Figure 3-5. The Create From Selection feature (discussed in Chapter 2) is used in the examples to name columns F through O with their row 10 headings.

| | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | | | | | |
| 2 | | | Outside | | | | | | | | | | |
| 3 | MEAN−2SIGMA | 45.72702047 | 13 | | | | | | | | | | |
| 4 | MEAN+2SIGMA | 126.8337101 | 35 | | | | | | | | | | |
| 5 | TOTAL | 1095 | | | | | | | | | | | |
| 6 | WITHIN 2 SIGMA | 1047 | | | | | | | | | | | |
| 7 | FRACTION | 95.62% | | | | | | | | | | | |
| 8 | | | | | | | | | | | | | |
| 9 | | | | | | | | | | | | | |
| 10 | weekday | Date | Cakes | Pies | Cookies | Smoothies | Coffee | promotion | month | Namemonth | | | |
| 11 | 2 | 1/1/2013 | 79 | 46 | 518 | 60 | 233 | none | 1 | January | | | |
| 12 | 3 | 1/2/2013 | 91 | 50 | 539 | 161 | 427 | none | 1 | January | | | |
| 13 | 4 | 1/3/2013 | 47 | 60 | 222 | 166 | 347 | none | 1 | January | | How many promotions? | |
| 14 | 5 | 1/4/2013 | 89 | 64 | 734 | 153 | 358 | none | 1 | January | | Promotion | 110 |
| 15 | 6 | 1/5/2013 | 112 | 73 | 764 | 240 | 392 | none | 1 | January | | None | 985 |
| 16 | 7 | 1/6/2013 | 89 | 57 | 922 | 259 | 510 | none | 1 | January | | | |
| 17 | 1 | 1/7/2013 | 70 | 50 | 476 | 120 | 334 | none | 1 | January | | Days sold>=122 cakes | 56 |

**Figure 3-5:** Illustrating the Rule of Thumb

## Counting the Number of Promotions

The Excel `COUNTIF` function can be used to count the number of rows in a range of cells that meet a single criterion. For example, you can compute the number of days on which La Petit Bakery had a promotion and did not have a promotion by using the `COUNTIF` function. Simply copy the formula `=COUNTIF(promotion,Q14)` from R14 to R15. You will see that there were 110 days with a promotion and 985 without a promotion. Note that Q14 as a criterion ensures that you only count days on which there was a promotion and Q15 as a criterion ensures that you only count days on which there was no promotion.

## Verifying the Rule of Thumb

In the "Rule of Thumb" section of this chapter you learned that the rule of thumb tells you that for approximately 95 percent of all days, cake sales should be within 2s of the mean. You can use the `COUNTIF` function to check if this is the case. In cell G3 you can compute for daily cake sales the Average –2s with the formula `AVERAGE(Cakes)−2*STDEV(Cakes)`. In G4 you can compute the Average +2s with the formula `AVERAGE(Cakes)+2*STDEV(Cakes)`.

You can also find the number of days (13) on which cakes were more than 2s below the mean. To do so, enter in cell H3 the formula `COUNTIF(Cakes,"<"&G3)`. The & (or concatenate) sign combines with the less than sign (in quotes because < is text) to ensure you only count entries in the Cakes column that are more than 2s (s = sample standard deviation) below the mean.

In cell H4 the formula `COUNTIF(Cakes,">"&G4)` computes the number of days (35) on which cake sales were more than 2s above the mean. Any data point that differs from the dataset's mean by more than 2s is called an `outlier`. Chapters 10 and 11 will make extensive use of the concept of an outlier.

The `COUNT` function can be used to count the number of numeric entries in a range while the `COUNTA` function can be used to count the number of non-blanks in a range. Also the `COUNTBLANK` function counts the number of blank cells in a range. For example, the function `=COUNT(Cakes)` in cell G5 counts how many numbers appear in the Cake column (1095). The `COUNTA` function counts the number of nonblank (text or numbers) cells in a range, and the `COUNTBLANK` function counts the number of blanks in a range. In G6 compute (with formula `=G5–H3–H4`) the number of cells (1047) within 2s of the mean. In G7 you find that 95.62 percent of the days had cake cells within 2s of the mean. This is in close agreement with the rule of thumb.

### Computing Average Daily Cake Sales

You can also use `SUMIF` and `COUNTIF` functions to calculate the average sales of cakes for each day of the week. To do so, perform the following steps:

1. Copy the formula `=SUMIF(daywk,Q20,Cakes)` from S20 to S21:S26 to compute the total cake sales for each day of the week.
2. Copy the formula `=COUNTIF(daywk,Q20)` from R20 to R21:R:26 to compute for each day of the week the number of times the day of the week occurs in your data set.
3. Finally, in Column T, divide the `SUMIF` result for each day by the `COUNTIF` result, and this gives the average cake sales for each day of the week.

To rank the average cake sales for each day of the week, copy the formula `=RANK(T20,$T$20:$T$26,0)` from U20 to U21:U26. The formula in U20 ranks the Monday average sales among all days of the week. Dollar signing the range T20:T26 ensures that when you copy the formula in U21, each day's average sales is ranked against all 7 days of the week. If you had not dollar signed the range T20:T26, then in U21 the range would have changed to T21:T27, which would be incorrect because this range excludes Monday's average sales. The last argument of 0 ensures that the day of the week with the largest sales is given a rank of 1. If you use a last argument of 1, then sales are ranked on a basis that makes the smallest number have a rank of 1, which is not appropriate in this situation. Note that Saturday is the best day for sales and Wednesday was the worst day for sales.

In an identical manner, as shown in Figure 3-6 you can compute the average daily sales of smoothies during each month of the year in the cell range Q28:U40. The results show that the summer months (June–August) are the best for smoothie sales.

| ◢ | Q | R | S | T | U |
|---|---|---|---|---|---|
| 18 | **Cakes** | | | | |
| 19 | Total Sales by day of week | How many days | Total Sales | Average Daily sales | Rank |
| 20 | Monday | 156 | 12259 | 78.58333 | 4 |
| 21 | Tuesday | 157 | 12079 | 76.93631 | 6 |
| 22 | Wednesday | 157 | 11986 | 76.34395 | 7 |
| 23 | Thursday | 157 | 12148 | 77.3758 | 5 |
| 24 | Friday | 156 | 14770 | 94.67949 | 3 |
| 25 | Saturday | 156 | 16181 | 103.7244 | 1 |
| 26 | Sunday | 156 | 15054 | 96.5 | 2 |
| 27 | **Smoothies** | | | | |
| 28 | Total sales by month of year | How many days | Total Sales | Average Daily Sales | Rank |
| 29 | January | 93 | 15955 | 171.5591 | 12 |
| 30 | February | 84 | 14692 | 174.9048 | 10 |
| 31 | March | 93 | 16930 | 182.043 | 8 |
| 32 | April | 90 | 19016 | 211.2889 | 7 |
| 33 | May | 93 | 22434 | 241.2258 | 4 |
| 34 | June | 90 | 24971 | 277.4556 | 3 |
| 35 | July | 93 | 26654 | 286.6022 | 2 |
| 36 | August | 93 | 26724 | 287.3548 | 1 |
| 37 | September | 90 | 19737 | 219.3 | 5 |
| 38 | October | 93 | 19813 | 213.043 | 6 |
| 39 | November | 90 | 15463 | 171.8111 | 11 |
| 40 | December | 93 | 16832 | 180.9892 | 9 |

**Figure 3-6:** Daily summary of cake sales

In the next section you will learn how the AVERAGEIF function makes these calculations a bit simpler.

## Using the COUNTIFS, SUMIFS, AVERAGEIF, and AVERAGEIFS Functions

COUNTIF and SUMIF functions are great, but they are limited to calculations based on a single criteria. The next few examples involve the use of four Excel functions that were first introduced in Excel 2007: COUNTIFS, SUMIFS, AVERAGEIF, and AVERAGEIFS (see the New 2007 Functions worksheet). These functions enable you to do calculations involving multiple (up to 127!) criteria. A brief description of the syntax of these functions follows:

■ The syntax of COUNTIFS is COUNTIFS(range1,criteria1, range2,critieria2, .. range_n, criteria_n). COUNTIFS counts the number of rows in which

the *range1* entry meets *criteria1*, the *range2* entry meets *criteria2*, and … the *range_n* entry meets *criteria_n*.

- The syntax of SUMIFS is SUMIFS(sum_range, range1, criteria1, range2, criteria2, ...,range_n,criteria_n). SUMIFS sums up every entry in the sum_range for which *criteria1* (based on *range1*), *criteria2* (based on *range2*), … *criteria_n* (based on *range_n*) are all met.

- The AVERAGEIF function has the syntax AVERAGEIF(range, criteria, average_range). AVERAGEIF averages the range of cells in the average range for which the entry in the range column meets the criteria.

- The syntax of AVERAGEIFS is AVERAGEIFS(average_range, criteria_range1, criteria_range2, ..., criteria_range_n). AVERAGEIFS averages every entry in the average range for which *criteria1* (based on *range1*), *criteria2* (based on *range2*), … *criteria_n* (based on *range_n*) are all met.

Now you can use these powerful functions to perform many important computations. The following examples are shown in the New 2007 Functions worksheet. The work is shown in Figure 3-7.

| | Q | R | S | T | U | V |
|---|---|---|---|---|---|---|
| 1 | Promotions on each day of week | | | Average Cookies Sales by day of week | | |
| 2 | COUNTIFS | How Many | | AVERAGEIF | Mean sales | |
| 3 | Monday | 23 | | Monday | 482.0962 | |
| 4 | Tuesday | 14 | | Tuesday | 480.6242 | |
| 5 | Wednesday | 17 | | Wednesday | 486.8662 | |
| 6 | Thursday | 13 | | Thursday | 488.6752 | |
| 7 | Friday | 13 | | Friday | 597.3397 | |
| 8 | Saturday | 14 | | Saturday | 654.7179 | |
| 9 | Sunday | 16 | | Sunday | 593.8974 | |
| 10 | | | | | | |
| 11 | Total Sales | | | | | |
| 12 | SUMIFS | | | | | |
| 13 | | Cakes | Pies | Cookies | Smoothie | Coffee |
| 14 | January | 7726 | 4803 | 48070 | 15955 | 37039 |
| 15 | February | 7089 | 4507 | 44886 | 14692 | 33514 |
| 16 | March | 7848 | 4999 | 50764 | 16930 | 36317 |
| 17 | April | 7601 | 4647 | 46740 | 19016 | 34550 |
| 18 | May | 8360 | 4846 | 49865 | 22434 | 34396 |
| 19 | June | 7760 | 4948 | 47298 | 24971 | 34938 |
| 20 | July | 8085 | 4881 | 50228 | 26654 | 37015 |
| 21 | August | 7898 | 5130 | 50674 | 26724 | 37784 |
| 22 | September | 7842 | 4712 | 49605 | 19737 | 34081 |
| 23 | October | 8041 | 5117 | 51527 | 19813 | 36518 |
| 24 | November | 8045 | 4955 | 51617 | 15463 | 35713 |
| 25 | December | 8182 | 5090 | 50520 | 16832 | 36966 |

**Figure 3-7:** Monthly sales summary

### Calculating the Number of Promotions on Each Day of the Week

In R3 the formula `COUNTIFS(daywk,Q3,promotion,"promotion")` counts the number of promotions (23) on Monday. Copying this formula to R4:R9 calculates the number of promotions on each day of the week.

### Calculating the Average Cookie Sales on Each Day of the Week

The formula in U3 `=AVERAGEIF(daywk,T3,Cookies)` averages the number of cookie sales on Monday. Copying this formula to the range U4:U9 calculates average cookie sales for all other days of the week.

### Computing Monthly Sales for each Product

The formula `=SUMIFS(INDIRECT(R$13),Namemonth,$Q14)` in cell R14 computes total cake sales in January (7726).

1. Place the `INDIRECT` function before the reference to R13 to enable Excel to recognize the word "Cakes" as a range name.
2. Copy this formula to the range R14:V25. This calculates the total sales for each product during each month.

> **NOTE**    Note the `INDIRECT` function makes it easy to copy formulas involving range names. If you do not use the `INDIRECT` function then Excel will not recognize text in a cell such as R13 as a range name.

### Computing Average Daily Sales by Month for Each Product

The formula `=AVERAGEIFS(INDIRECT(R$28),Namemonth,$Q29)` in cell R29 computes average cake sales in January. Copying this formula to the range R29:V40 calculates average sales during each month for each product. Figure 3-8 shows this example.

## Summarizing Data with Subtotals

The final method to summarize market data discussed in this chapter is with *subtotals*. The subtotals feature yields a great looking summary of data. Unfortunately, if new data is added, subtotals are difficult to update. Suppose you want to get a breakdown for each day of the week and for each product showing how sales differ on days with and without promotions (see the `Subtotals Bakery` worksheet). Before computing subtotals you need to sort your data in the order in which you want your subtotals to be computed. In the subtotals you want to see the day of the week, and

then "promotion" or "not promotion." Therefore, begin by sorting the data in this order. The dialog box to create this sort is shown in Figure 3-9.

| | Q | R | S | T | U | V |
|---|---|---|---|---|---|---|
| 27 | **Average Sales** | | | | | |
| 28 | **AVERAGEIFS** | Cakes | Pies | Cookies | Smoothie | Coffee |
| 29 | January | 83.0752688 | 51.64516 | 516.8817204 | 171.5591 | 398.2688 |
| 30 | February | 84.3928571 | 53.65476 | 534.3571429 | 174.9048 | 398.9762 |
| 31 | March | 84.3870968 | 53.75269 | 545.8494624 | 182.043 | 390.5054 |
| 32 | April | 84.4555556 | 51.63333 | 519.3333333 | 211.2889 | 383.8889 |
| 33 | May | 89.8924731 | 52.10753 | 536.1827957 | 241.2258 | 369.8495 |
| 34 | June | 86.2222222 | 54.97778 | 525.5333333 | 277.4556 | 388.2 |
| 35 | July | 86.9354839 | 52.48387 | 540.0860215 | 286.6022 | 398.0108 |
| 36 | August | 84.9247312 | 55.16129 | 544.8817204 | 287.3548 | 406.2796 |
| 37 | September | 87.1333333 | 52.35556 | 551.1666667 | 219.3 | 378.6778 |
| 38 | October | 86.4623656 | 55.02151 | 554.0537634 | 213.043 | 392.6667 |
| 39 | November | 89.3888889 | 55.05556 | 573.5222222 | 171.8111 | 396.8111 |
| 40 | December | 87.9784946 | 54.73118 | 543.2258065 | 180.9892 | 397.4839 |

**Figure 3-8:** Using AVERAGEIFS to summarize monthly sales
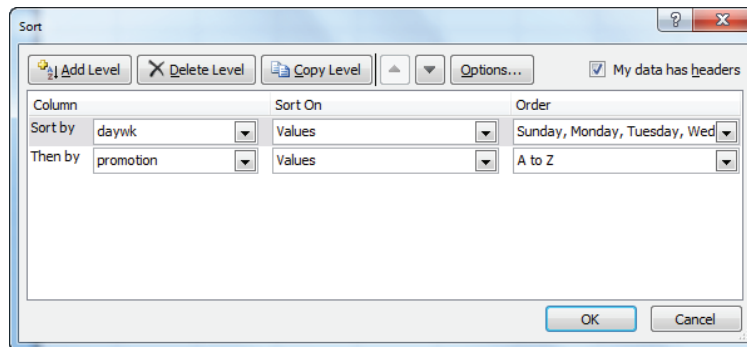


**Figure 3-9:** Sorting in preparation for subtotals

After performing the sort you see all the Sundays with no promotion, followed by the Sundays with a promotion, then the Mondays with no promotion days, and so on. Perform the following steps to continue:

1. Place your cursor anywhere in the data and select Subtotal from the Data tab on the ribbon.
2. Compute Subtotals for each product for each day of the week by filling in the Subtotals dialog box, as shown in Figure 3-10.
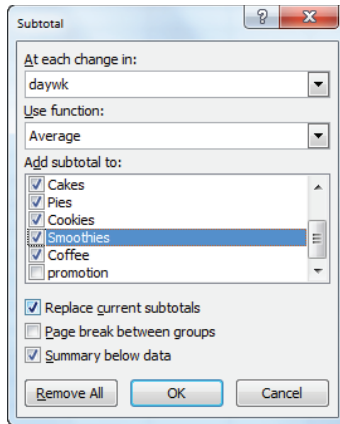
**Figure 3-10:** Subtotals settings for daily summary

3. Next you "nest" these subtotals with totals each day of the week for no promotion and promotion days. To compute the nested subtotals giving the breakdown of average product sales for each day of the week on promotion and no promotion days, select Subtotal from the Data tab and fill in the dialog box, as shown in Figure 3-11.

4. Select OK and the Subtotals feature creates for each day of the week total sales for each product. This summary (as shown in Figure 3-12) can be seen by clicking the number 2 in the upper left corner of your screen.
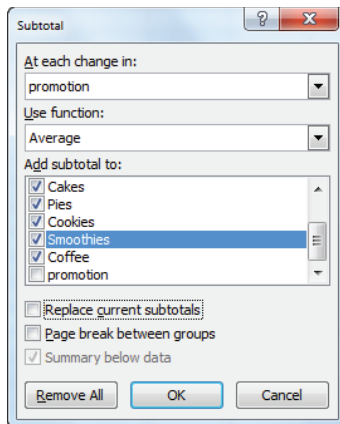


**Figure 3-11:** Final subtotals dialog box

By unchecking Replace Current Subtotals, you ensure that the subtotals on Promotions will build on and not replace the daily calculations. By clicking the 3 in the upper-left corner of the screen, you can find the final breakdown (see Figure

3-13) of average sales on no promotion versus promotion sales for each day of the week. It is comforting to note that for each product and day of week combination, average sales are higher with the promotion than without.

| | daywk | weekday | Date | Cakes | Pies | Cookies | Smoothies | Coffee | promotion |
|---|---|---|---|---|---|---|---|---|---|
| 9 | | | | | | | | | |
| 10 | daywk | weekday | Date | Cakes | Pies | Cookies | Smoothies | Coffee | promotion |
| 169 | Sunday Average | | | 96.5 | 59.46795 | 593.8974 | 240.2820513 | 434.5128 | |
| 328 | Monday Average | | | 78.58333 | 48.19231 | 482.0962 | 195.3653846 | 353.7564 | |
| 488 | Tuesday Average | | | 76.93631 | 46.76433 | 480.6242 | 198.5732484 | 348.3631 | |
| 648 | Wednesday Average | | | 76.34395 | 47.75159 | 486.8662 | 198.2101911 | 347.5287 | |
| 808 | Thursday Average | | | 77.3758 | 48.10828 | 488.6752 | 197.1719745 | 362.9427 | |
| 967 | Friday Average | | | 94.67949 | 59.6859 | 597.3397 | 235.0064103 | 427.0321 | |
| 1126 | Saturday Average | | | 103.7244 | 64.98077 | 654.7179 | 265.0512821 | 467.9936 | |
| 1127 | Grand Average | | | 86.28037 | 53.54795 | 540.4511 | 218.4666667 | 391.6265 | |

**Figure 3-12:** Daily subtotals summary

| | daywk | weekday | Date | Cakes | Pies | Cookies | Smoothies | Coffee | promotion | |
|---|---|---|---|---|---|---|---|---|---|---|
| 9 | | | | | | | | | | |
| 10 | daywk | weekday | Date | Cakes | Pies | Cookies | Smoothies | Coffee | promotion | |
| 151 | | | | 95.37857 | 58.67857 | 584.7571 | 236.0428571 | 428.4786 | none Average | |
| 168 | | | | 106.3125 | 66.375 | 673.875 | 277.375 | 487.3125 | promotion Average | |
| 169 | Sunday Average | | | 96.5 | 59.46795 | 593.8974 | 240.2820513 | 434.5128 | | |
| 303 | | | | 76.04511 | 47.30827 | 469.015 | 194.4135338 | 342.6241 | none Average | |
| 327 | | | | 93.26087 | 53.30435 | 557.7391 | 200.8695652 | 418.1304 | promotion Average | |
| 328 | Monday Average | | | 78.58333 | 48.19231 | 482.0962 | 195.3653846 | 353.7564 | | |
| 472 | | | | 76.13986 | 46.33566 | 469.9441 | 194.8391608 | 346.014 | none Average | |
| 487 | | | | 85.07143 | 51.14286 | 589.7143 | 236.7142857 | 372.3571 | promotion Average | |
| 488 | Tuesday Average | | | 76.93631 | 46.76433 | 480.6242 | 198.5732484 | 348.3631 | | |
| 629 | | | | 75.02143 | 47.06429 | 476.9786 | 195.75 | 343.2643 | none Average | |
| 647 | | | | 87.23529 | 53.41176 | 568.2941 | 218.4705882 | 382.6471 | promotion Average | |
| 648 | Wednesday Average | | | 76.34395 | 47.75159 | 486.8662 | 198.2101911 | 347.5287 | | |
| 793 | | | | 76.56944 | 47.34722 | 478.25 | 194.1597222 | 360.9722 | none Average | |
| 807 | | | | 86.30769 | 56.53846 | 604.1538 | 230.5384615 | 384.7692 | promotion Average | |
| 808 | Thursday Average | | | 77.3758 | 48.10828 | 488.6752 | 197.1719745 | 362.9427 | | |
| 952 | | | | 93.96503 | 59.07692 | 589.8392 | 227.2377622 | 421.8112 | none Average | |
| 966 | | | | 102.5385 | 66.38462 | 679.8462 | 320.4615385 | 484.4615 | promotion Average | |
| 967 | Friday Average | | | 94.67949 | 59.6859 | 597.3397 | 235.0064103 | 427.0321 | | |
| 1110 | | | | 102.1479 | 63.78169 | 650.7183 | 258.028169 | 459.5775 | none Average | |
| 1125 | | | | 119.7143 | 77.14286 | 695.2857 | 336.2857143 | 553.3571 | promotion Average | |
| 1126 | Saturday Average | | | 103.7244 | 64.98077 | 654.7179 | 265.0512821 | 467.9936 | | |
| 1127 | Grand Average | | | 86.28037 | 53.54795 | 540.4511 | 218.4666667 | 391.6265 | | |

**Figure 3-13:** Final subtotals summary

## USING EXCEL OUTLINES WITH THE SUBTOTALS FEATURE

Whenever you use the Subtotals feature you will see numbers (in this case 1,2, 3, and 4) in the upper left hand corner of your spreadsheet. This is an example of an Excel outline. The higher the number in the outline, the less aggregated is the data. Thus selecting 4 gives the original data (with subtotals below the relevant data), selecting 3 gives a sales breakdown by day of week and promotion and lack thereof, selecting 2 gives a sales breakdown by day of the week, and selecting 1 gives an overall breakdown of average sales by product. If your cursor is within the original data you can click the Remove All button in the dialog box to remove all subtotals.

## Using Array Formulas to Summarize *ESPN The Magazine* Subscriber Demographics

The six wonderful functions previously discussed are great ways to calculate conditional counts, sums, and averages. Sometimes, however, you might want to compute a conditional median, standard deviation, percentile, or some other statistical function. Writing your own array formulas you can easily create your own version of a MEDIANIF, STDEVIF, or other conditional statistical functions.

> **NOTE**  The rules for array formulas discussed in "Summarizing Data with a Histogram" also apply to array formulas that you might write.

To illustrate the idea, look at the Chapter 1 data on ESPN subscribers located in the Slicing with arrays worksheet (see also Figure 3-14).

| | A | B | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | | | |
| 2 | | | | | | | | | | | |
| 3 | | | | | | | | | | | |
| 4 | | | | | | | | | | | |
| 5 | | | | | | | | | | | |
| 6 | | | | | | | | | | | |
| 7 | | | | | | | | | | | |
| 8 | Age | Gender | Income | Location | | | | | | | |
| 9 | 72 | m | 72 | rural | | 77 | Median income subscibers >= 50 | | | | |
| 10 | 29 | m | 68 | suburban | | 80 | Median incomesubscribers <50 | | | | |
| 11 | 33 | m | 57 | suburban | | 27.36277 | Standard deviation income rural subscribers | | | | |
| 12 | 25 | m | 62 | suburban | | | | | | | |
| 13 | 38 | m | 164 | urban | | | | | | | |
| 14 | 33 | m | 44 | urban | | | | | | | |
| 15 | 18 | f | 62 | rural | | | | | | | |
| 16 | 17 | m | 68 | urban | | | | | | | |
| 17 | 32 | f | 53 | urban | | | | | | | |
| 18 | 24 | f | 92 | urban | | | | | | | |
| 19 | 26 | f | 54 | suburban | | | | | | | |
| 20 | 40 | m | 88 | urban | | | | | | | |
| 21 | 26 | f | 46 | rural | | | | | | | |
| 22 | 21 | m | 83 | suburban | | | | | | | |
| 23 | 29 | m | 144 | rural | | | | | | | |
| 24 | 51 | f | 30 | suburban | | | | | | | |

**Figure 3-14:** Using arrays to compute conditional medians and standard deviations

Suppose you have been asked to determine the median income for subscribers over 50. To accomplish this goal, perform the following steps:

1. Array enter in cell F9 the formula =MEDIAN(IF(Age>=50,Income,"")). This formula loops through the Income column and creates an array as follows: whenever the Age column contains a number greater than or equal to 50, the array returns the income value; otherwise the array returns a blank. Now the MEDIAN function is simply being applied to the rows corresponding to subscribers who are at least 50 years old. The median income for subscribers who are at least 50 is $77,000. Note that since your array formula populates a single cell you do not need to select a range containing more than one cell before typing in the formula.

2. In a similar fashion, array enter in cell F10 the formula =MEDIAN(IF(Age<50,Income,"")) to compute the median salary ($80,000) for subscribers who are under 50 years old.

3. Finally, in cell F11, array enter the formula =STDEV(IF(Location="rural", Income,"")) to compute (27.36) the standard deviation of the Income for Rural subscribers. This formula produces an array that contains only the Income values for rural subscribers and takes the standard deviation of the values in the new array.

# Summary

In this chapter you learned the following:

- Using the FREQUENCY array function and the TABLE feature, you can create a histogram that summarizes data and automatically updates to include new data.

- The median (computed with the MEDIAN function) is used to summarize a typical value for a highly skewed data set. Otherwise, the mean (computed with the AVERAGE function) is used to summarize a typical value from a data set.

- To measure a data set's spread about the mean, you can use either variance (computed with the Excel VAR function) or standard deviation (computed with the Excel STDEV function). Standard deviation is the usual measure because it is in the same units as the data.

- For data sets that do not exhibit significant skewness, approximately 95 percent of your data is within two standard deviations of the mean.
- The PERCENTILE.EXC function returns a given percentile for a data set.
- The PERCENTRANK.EXC function returns the percentile rank for a given value in a data set.
- The LARGE and SMALL functions enable you to compute either the *k*th largest or *k*th smallest value in a data set.
- The COUNTIF and COUNTIFS functions enable you to count how many rows in a range meet a single or multiple criteria.
- The SUMIF and SUMIFS enable you to sum a set of rows in a range that meets a single or multiple criteria.
- The AVERAGEIF and AVERAGEIFS functions enable you to average a set of rows in a range that meets a single or multiple criteria.
- The SUBTOTALS feature enables you to create an attractive summary of data that closely resembles a PivotTable. Data must be sorted before you invoke the SUBTOTALS feature.
- Using array functions you can create formulas that mimic a STDEVIF, MEDIANIF, or PERCENTILEIF function.

## Exercises

1. Exercises 1-6 use the data in the Descriptive stats worksheet of the Chapter3bakery.xlsx workbook. Create a dynamically updating histogram for daily cookie sales.
2. Are daily cookie sales symmetric or skewed?
3. Fill in the blank: By the rule of thumb you would expect on 95 percent of all days daily smoothie demand will be between ___ and ___.
4. Determine the fraction of days for which smoothie demand is an outlier.
5. Fill in the blank: There is a 10 percent chance daily demand for smoothies will exceed ____.
6. Fill in the blank: There is a____ chance that at least 600 cookies will be sold in a day.
7. Exercises 7 and 8 use the data in the Data worksheet of the ESPN.xlsx workbook. Find the average age of all ESPN subscribers who make at least $100,000 a year.
8. For each location and age group (under 25, 25–39, 40–54, 55, and over) determine the fraction of ESPN subscribers for each location that are in each age group.

9. Exercises 9-12 use the data in the `Descriptive stats` worksheet of the `Chapter3bakery.xlsx` workbook. Determine the percentage of La Petit Bakery cookie sales for each day of the week and month combination. For example, your final result should tell you the fraction of cookie sales that occur on a Monday in January, and so on.

10. Determine the average profit earned for each day of the week and month combination. Assume the profit earned by La Petit Bakery on each product is as follows:
    - Cakes: $2
    - Cookie: $0.50
    - Pie: $1.50
    - Smoothie: $1.00
    - Coffee: $0.80

11. Find the median cake sales on days in which at least 500 cookies were sold.

12. Fill in the blank: On days in which at least 500 cookies are sold there is a 5 percent chance at least _____ cookies are sold.

13. In the years 1980–2012 median U.S. family income (after inflation) has dropped but mean family income has sharply increased. Can you explain this seeming anomaly?

14. Compute and interpret the skewness for the three data sets in the file `Skewexamples.xlsx`.

# II

# Pricing

# 4  Estimating Demand Curves and Using Solver to Optimize Price

U nderstanding how pricing impacts revenues and profitability is one of the most important issues faced by managers. To do so, managers need to understand how consumers' willingness to purchase changes at different price levels and how these changes impact profitability. (That is, managers need to understand the demand curve.) To do so effectively, this chapter covers the following topics:

- Using the Excel Trend Curve and Goal Seek to obtain back-of-the-envelope estimation of a demand curve.
- Using the Excel Solver to determine the profit maximizing price.
- Examining the effect of product tie-ins on the optimal product price.
- Using the SolverTable add-in to quickly price thousands of products!

## Estimating Linear and Power Demand Curves

In this section you learn how to fit the two most frequently used demand curves (linear and power) to a particular marketing situation. These estimations are used to determine a profit-maximizing price, and to do so you need to know two things:

- The variable cost to produce each unit of the product. (Call this *UC*.)
- The product's demand curve. Simply put, a demand curve tells you the number of units of your product a customer will demand at each price. In short, if you charge a price of $\$p$ per unit, the demand curve gives you a number $D(p)$, which equals the number of units of your product that will be demanded at price $\$p$. Of course, a firm's demand curve is constantly changing and often depends on factors beyond the firm's control (such as the state of the economy and a competitor's price). Part III, "Forecasting," addresses these factors.

When you know UC and the demand curve, the profit corresponding to a price of $\$p$ is simply $(p - UC) * D(p)$. After you have an equation for $D(p)$, which gives the

quantity of the product demanded for each price, you can use the Microsoft Office Excel Solver feature to find the profit-maximizing price.

## Price Elasticity

Given a demand curve, the *price elasticity* for demand is the percentage of decrease in demand resulting from a 1 percent increase in price. When elasticity is larger than 1, demand is price elastic. When demand is price elastic, a price cut will increase revenue. When elasticity is less than 1, demand is price inelastic. When demand is price inelastic, a price cut will decrease revenue. Studies by economists have obtained the following estimates of elasticity:

- **Salt:** 0.1 (very inelastic)
- **Coffee:** 0.25 (inelastic)
- **Legal fees:** 0.4 (inelastic)
- **TV sets:** 1.2 (slightly elastic)
- **Restaurant meals:** 2.3 (elastic)
- **Foreign travel:** 4.0 (very elastic)

A 1 percent decrease in the cost of foreign travel, for example, can result in a 4 percent increase in demand for foreign travel. Managers need to understand the price elasticity at each price point to make optimal pricing decisions. In the next section you will use price elasticity to estimate a product's demand curve.

## Forms of Demand Curves

There are multiple forms of demand curves that you can use to analyze marketing data. Using $q$ to represent the quantity demanded of a product, the two most commonly used forms for estimating demand curves are as follows:

- **Linear demand curve:** In this case, demand follows a straight line relationship of the form $q = a - bp$. Here $q$ = quantity demanded and $p$ = unit price. For example, $q = 10 - p$ is a linear demand curve. (Here $a$ and $b$ can be determined by using a method described in the "Estimating a Linear Demand Curve" section of this chapter.) When the demand curve is linear, the elasticity is changing.
- **Power demand curve:** In this situation, the demand curve is described by a power curve of the form $q = ap^b$, $a > 0$, $b < 0$. Again, $a$ and $b$ can be determined by the method described later in the chapter. The equation $q = 100p^{-2}$ is an example of a power demand curve. If demand follows a power curve, for any price, the elasticity equals $-b$. See Exercise 11 for an explanation of this

important property of the power demand curve. Thus, for the demand curve $q = 100p^{-2}$ the price elasticity of demand always equals 2.

## Estimating a Linear Demand Curve

Suppose that a product's demand curve follows a linear demand curve. Given the current price and demand for a product and the product's price elasticity of demand, determining the product's demand curve is a simple matter. The following example illustrates how to fit a linear demand curve.

Suppose a product is currently selling for $100 and demand equals 500 units. The product's price elasticity for demand is 2. Assuming the demand curve is linear, you can determine the equation of the demand curve. The solution is in the `Linearfit` `.xls` file, which is shown in Figure 4-1.
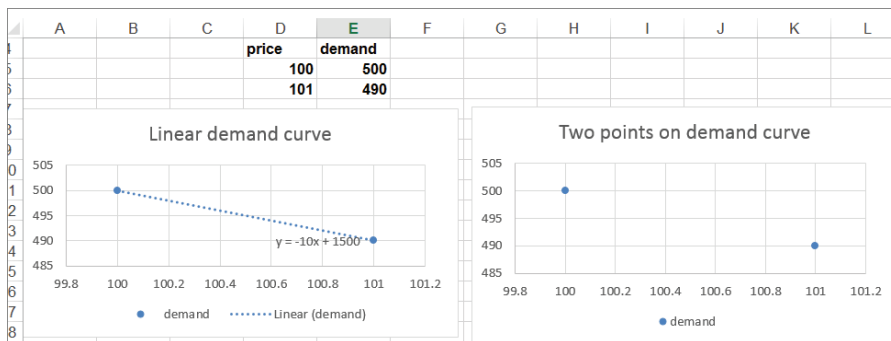


**Figure 4-1:** Fitting linear demand curve

Given two points, you know that there is a unique straight line that passes through those two points. You actually know two points on your demand curve. One point is $p$=100 and $q$=500. Because elasticity of demand equals 2, a 1 percent increase in price results in a 2 percent decrease in demand. Thus, if $p$=101 (a 1 percent increase), demand drops by 2 percent of 500 (10 units) to 490. Thus $p$=101 and $q$=490 is a second point on the demand curve. You can now use the Excel Trendline to find the straight line that passes through the points (100,500) and (101,490).

Begin by entering these points in the worksheet in the cell range D5:E6 (refer to Figure 4-1). Then select the range D4:E6, and on the Ribbon, in the Charts group choose Scatter ➤ Scatter with Only Markers.

1. Begin by entering the points in the `Linearfit.xls` worksheet in the cell range D5:E6 (refer to Figure 4-1).
2. Select the range D4:E6.

3. Go to the Insert tab and in the Charts group choose Scatter ➤ Scatter with Only Markers. You should now see that the graph has a positive slope. This would imply that higher prices lead to higher demand, which cannot be correct. The problem is that with only two data points, Excel assumes that the data points you want to graph are in separate columns, not separate rows.

4. To ensure Excel understands that the individual points are in separate rows, click inside the graph.

5. On the Ribbon click the Design tab in the Chart Tools section.

6. Click Switch Row/Column in the Data section of the Design tab.

**NOTE**   Note that by clicking the Select Data button, you can change the source data that generates your chart.

7. Next, right-click one of the points and then click Add Trendline.

8. Click the Linear button and click Display Equation on Chart option.

9. Click Close in the Format Trendline box.

You see the straight line plot, complete with the equation referred to in Figure 4-1. Because *x* is price and *y* is demand, the equation for your demand curve is $q = 1500 - 10p$. This equation means that each \$1 increase in price costs 10 units of demand. Of course, demand cannot be linear for all values of *p* because for large values of *p*, a linear demand curve yields negative demand. For prices near the current price, however, the linear demand curve is usually a good approximation to the product's true demand curve.

## Estimating a Power Demand Curve

Recall that for a linear demand curve the price elasticity is different for each price point. If the marketing analyst believes that elasticity remains relatively constant as price changes, then she can use a power demand curve (which has constant price elasticity) to model demand for a product.

Again assume that a product is currently selling for \$100 and demand equals 500 units. Assume also that the product's price elasticity for demand is known to equal 2. (In Chapters 31 and 32 you will learn some more advanced methods to estimate price elasticity). You can fit a power demand curve to this information by performing the following steps in the `Powerfit.xls` file, as shown in Figure 4-2.

|   | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | | | |
| 2 | | | | | a | | | | | |
| 3 | | | | | 5000000 | | | | | |
| 4 | | | | price | demand | | | Use goal seek | | |
| 5 | | | | 100 | 500 | | | Set cell E5 to 500 by changing | | |
| 6 | | | | | | | | e3 | | |
| 7 | | | | | | | Price | Demand | | |
| 8 | | | | | | | 50 | 2000 | | |
| 9 | | | | | | | 60 | 1388.889 | | |
| 10 | | | | | | | 70 | 1020.408 | | |
| 11 | | | | | | | 80 | 781.25 | | |
| 12 | | | | | | | 90 | 617.284 | | |
| 13 | | | | | | | 100 | 500 | | |
| 14 | | | | | | | 110 | 413.2231 | | |
| 15 | | | | | | | 120 | 347.2222 | | |
| 16 | | | | | | | 130 | 295.858 | | |
| 17 | | | | | | | 140 | 255.102 | | |
| 18 | | | | | | | 150 | 222.2222 | | |
| 19 | | | | | | | | | | |
| 20 | | | | | | | | | | |

Chart: Power demand curve, $y = 5000000.000x^{-2.000}$, with Demand and Power (Demand) series, price on x-axis, demand on y-axis.
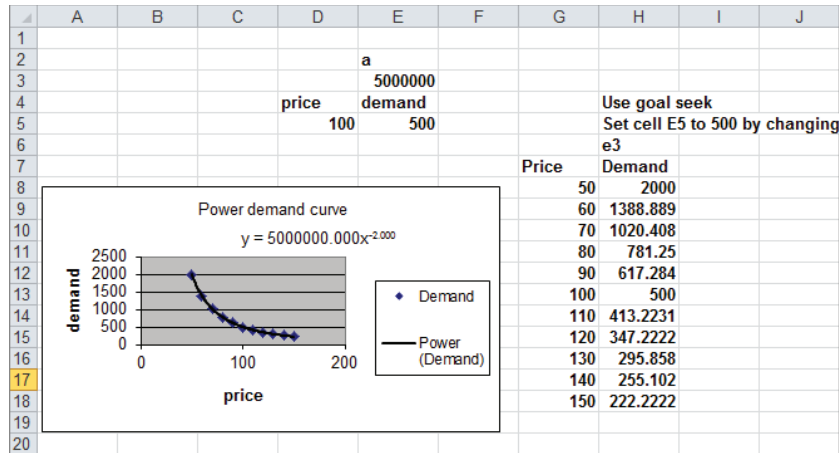
**Figure 4-2:** Fitting Power Demand Curve

1. After naming cell E3 as *a*, enter a trial value for *a*.

2. In cell D5, enter the current price of $100. Because elasticity of demand equals 2, you know that the demand curve has the form $q=ap^{-2}$ where *a* is unknown.

3. In cell E5, enter a formula that computes the demand when the unit price equals $100. The demand is linked to your choice of the value of *a* in cell E3 with the formula a*D5^−2.

4. Now use the Global Seek command to determine the value of *a*. This formula makes the demand for price $100 equal to 500 units. Goal Seek enables you to find a value of a cell in a spreadsheet (called the By Changing Cell) that makes a formula (called the Set Cell) hit a wanted Value. In the example you want to change cell E3 so that the formula in cell E5 equals 500. Set cell E5 to the value of 500 by changing cell E3.

5. To invoke Goal Seek, switch to the Data tab and select What-If analysis from the Data Tools Group, then choose Goal Seek from the drop down arrow. Fill the dialog box as shown in Figure 4-3. With these settings, Excel changes the changing cell (E3) until the value of the set cell (E5) matches your desired value of 500.

A value for *a* = 5 million yields a demand of 500 at a price of $100. Thus, the demand curve (refer to Figure 4-2) is given by $q = 5,000,000p^{-2}$. For any price, the price elasticity of demand on this demand curve equals 2.
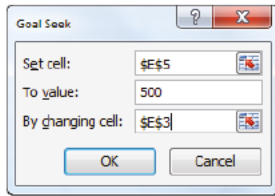
**Figure 4-3:** Goal Seek dialog box

## Using the Excel Solver to Optimize Price

Often in marketing you want to maximize profit, minimize, or optimize some other objective. The *Excel Solver* is a powerful tool that you can use to solve many marketing (and other!) optimization problems. The Solver is an add-in. To activate the Solver proceed as follows:

### For Excel 2010 or Excel 2013:

1. Select File and then Options.
2. Select Add-Ins, click Go, check the Solver add-in, and select OK.
3. Now click the Data tab and you should see the Solver add-in on the right side of the Ribbon.

### For Excel 2007:

1. Click the Office Button (the oval in the left side of the ribbon) and choose Excel Options.
2. Selecting Add-Ins, click Go, check the Solver add-in, and select OK.
3. Now click the Data tab and you should see the Solver add-in on the right side of the Ribbon.

In this chapter the examples work with the Excel 2010 Solver, which is more powerful than the Solver included with previous versions of Excel. If you select Solver from the Data tab, the Solver window appears, as shown in Figure 4-4.

To define a Solver model, you must specify in the Solver dialog box the following three parts of the model:

- Set Objective or Target Cell: The objective cell contains the goal you want to maximize (like profit) or minimize (production cost).

■ **Changing Variable Cells:** These cells are the cells that you can change or adjust to optimize the target cell. In this chapter the changing cells will be each product's price.

■ **Constraints:** These are restrictions on the changing cells. For example, you might want to constrain the price for each of your products to be within 10 percent of the competitor's price.
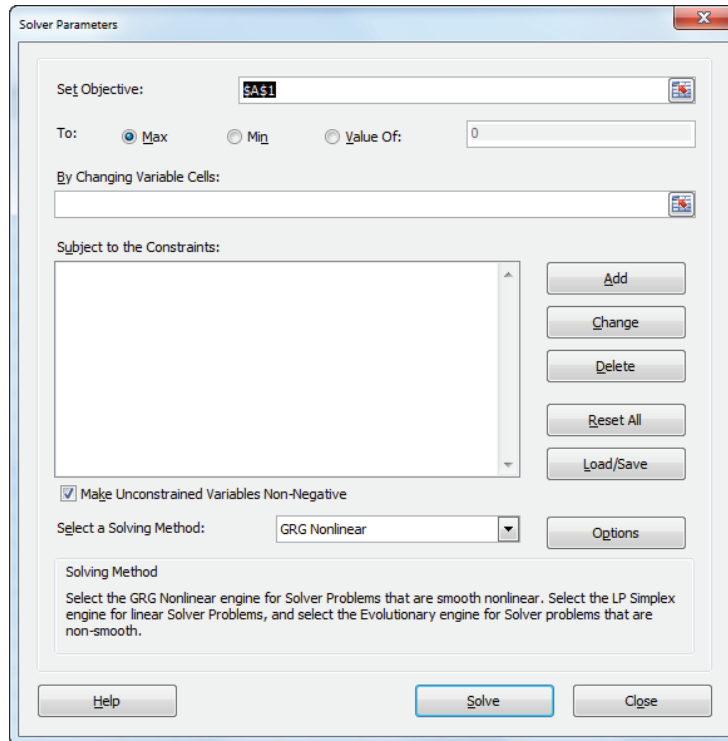


**Figure 4-4:** Solver window

The Excel Solver has been greatly revamped and improved in Excel 2010. The primary change is the presence of the Select a Solving Method drop-down list. From this list you must select the appropriate solution engine for your optimization problem. You can choose from the following options:

■ The *Simplex LP engine* is used to solve linear optimization problems. A linear optimization problem is one in which the target cell and constraints are all created by adding together terms of the form *(changing cell)\*(constant)*. Most

marketing models are not linear. An exception is the classic advertising media selection model discussed in Chapter 35, "Media Selection Models."

- The *GRG Nonlinear engine* is used to solve optimization problems in which the target cell and/or some of the constraints are not linear and are computed by using typical mathematical operations such as multiplying or dividing changing cells, raising changing cells to a power, using exponential or trig functions involving changing cells, and so on. The GRG engine includes a powerful Multistart option that enables users to solve many problems that were solved incorrectly with previous versions of Excel. The Multistart option will be used extensively throughout this book.

- The *Evolutionary engine* is used when your target cell and/or constraints contain nonsmooth functions that reference changing cells. For example, if your target cell and/or constraints contain `IF`, `SUMIF`, `COUNTIF`, `SUMIFS`, `COUNTIFS`, `AVERAGEIF`, `AVERAGEIFS`, `ABS`, `MAX`, or `MIN` functions that reference the changing cells, then the Evolutionary engine probably has the best shot at finding a good solution to your optimization problem. The Evolutionary engine is extensively used throughout this book.

After you have input the target cell, changing cells, and constraints, what does Solver do? A set of values for the changing cells is a *feasible solution* if it meets all constraints, and the Solver essentially searches through all feasible solutions and finds the set of feasible solution changing cell values (called the *optimal solution*) that has the best value for the target cell (largest in a maximization and smallest in a minimization). If there is more than one optimal solution, the Solver stops at the first one it finds.

## Pricing Razors (No Blades!)

Using the techniques described in the "Estimating a Linear Demand Curve" section, it's easy to determine a demand curve for the product that's originally purchased. You can then use the Microsoft Office Excel Solver to determine the original product price that maximizes the sum of the profit earned from razors. Then you can show how the fact that purchasers of razors also buy blades reduces the profit maximizing price for razors.

Suppose that you currently charge $5.00 for a razor and you sell 6 million razors. Assume that the variable cost to produce a razor is $2.00. Finally, suppose that the

price elasticity of demand for razors is 2, and the demand curve is linear. What price should you charge for razors?

You can determine a demand curve (assuming a linear demand curve), as shown in Figure 4-5. (You can find this data and the chart on the `no blades` worksheet in the file `razorsandblades.xls`.) Two points on the demand curve are price = $5.00, demand = 6 million razors and price = $5.05 (an increase of 1 percent), demand = 5.88 million (2 percent less than 6 million).

1. Begin by drawing a chart and inserting a linear trend line, as shown in the section "Estimating a Linear Demand Curve." You'll find the demand curve equation is $y = 18 - 2.4x$. Because $x$ equals price and $y$ equals demand, you can write the demand curve for razors as follows:

$$\text{Demand (in millions)} = 18 - 2.4(\text{price})$$
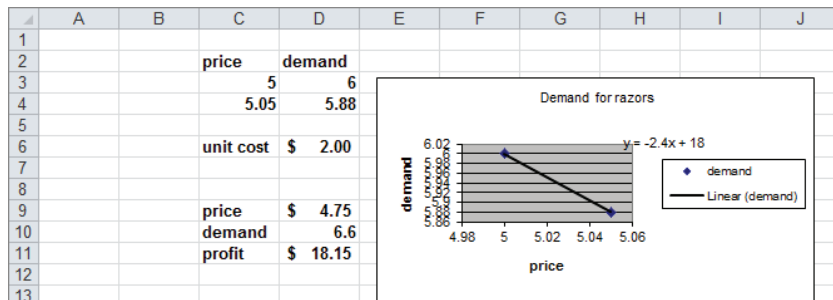


**Figure 4-5:** Optimizing razor price: no blades

2. Associate the names in cell C6 and the range C9:C11 with cells D6 and D9:D11.

3. Next, enter a trial price in D9 to determine demand for that price in cell D10 with the formula `=18-2.4*price`.

4. Determine in cell D11 the profit for razors by using the formula `=demand*(price-unit_cost)`.

5. You can now use Solver to determine the profit-maximizing price. The Solver Parameters dialog box is shown in Figure 4-6. Choose to maximize the profit cell (cell D11) by changing the price (cell D9).

6. The model is not linear because the target cell multiplies together two quantities—*demand* and *(price–cost)*—each depending on the changing cell. Therefore choose the GRG Nonlinear option. Solver finds that by charging $4.75 for a razor, you can maximize the profit. (The maximum profit is $18.15 million.)
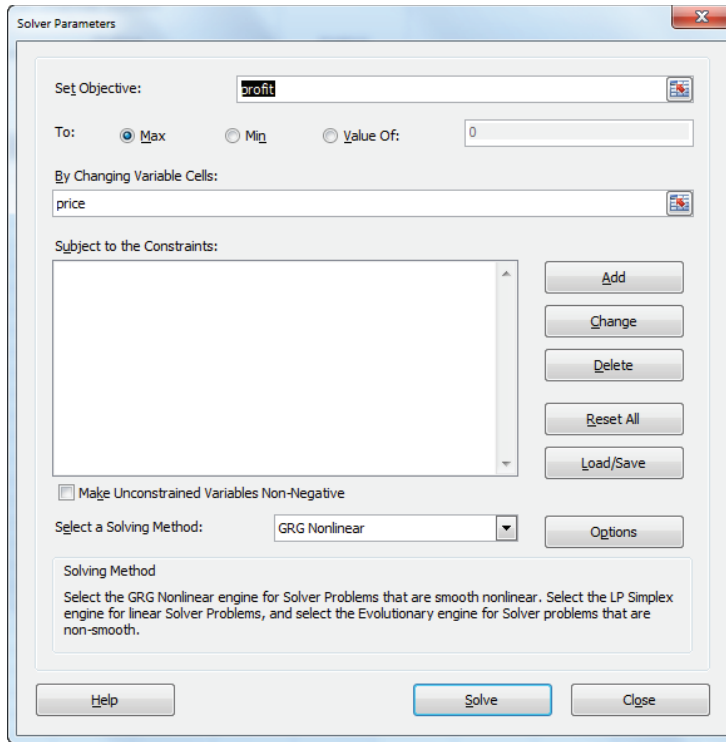
**Figure 4-6:** Solver window for razor price: no blades

## Incorporating Complementary Products

Certain consumer product purchases frequently result in the purchase of related products, or *complementary products*. Table 4-1 provides some examples:

**Table 4-1:** Examples of Complementary Product

| Original Purchase | Tie-in Complementary Product |
| --- | --- |
| Men's suit | Tie or shirt |
| Inkjet printer | Printer cartridge |
| Xbox console | Video game |
| Cell phone | Case |

If the profit from complementary products is included in the target cell, the profit maximizing price for the original product will decrease. Suppose that the average purchaser of a razor buys 50 blades and you earn $0.15 of profit per blade

purchased. You can use the Excel Solver to determine how this changes the price you should charge for a razor. Assume that the price of a blade is fixed. (In Exercise 3 at the end of the chapter, the blade price changes.) The analysis is in the `blades` worksheet of the `razorsandblades.xls` file, as shown in Figure 4-7.
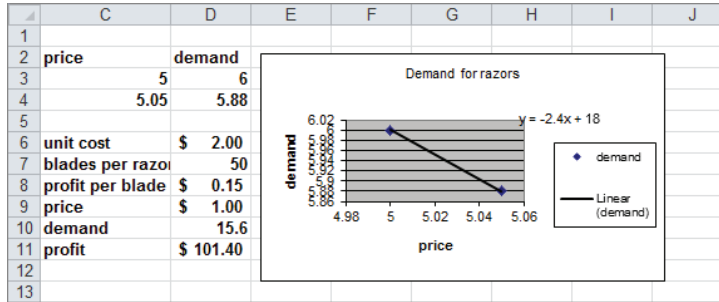


**Figure 4-7:** Optimizing razor price: blade profit included

To maximize profit perform the following steps:

1. Use the Create from Selection command in the Defined Names group on the Formulas tab to associate the names in cells C6:C11 with cells D6:D11. (For example, cell D10 is named Demand.)

   **NOTE**    Recall that cell D10 of the `no blades` worksheet is also named Demand. When you use the range name Demand in a formula Excel simply refers to the cell named Demand in the current worksheet. In other words, when you use the range name Demand in the `blades` worksheet, Excel refers to cell D10 of that worksheet, and not to cell D10 in the `no blades` worksheet.

2. In cells D7 and D8, enter the relevant information about blades.
3. In D9, enter a trial price for razors, and in D10, compute the demand with the formula `18-2.4*price`.
4. Next, in cell D11, compute the total profit from razors and blades with the following formula:

   `demand*(price-unit_cost)+demand*blades_per_razor*profit_per_blade`

   Here the `demand*blades_per_razor*profit_per_blade` is the profit from blades.

5. The Solver setup is exactly as shown earlier in Figure 4-6: Change the price to maximize the profit. Of course, now the profit formula includes the profit earned from blades.

Profit is maximized by charging only $1.00 (half the variable cost!) for a razor. This price results from making so much money from the blades. You are much better off ensuring that many people have razors even though you lose $1.00 on each razor sold. Many companies do not understand the importance of the profit from tie-in products. This leads them to overprice their primary product and not maximize their total profit.

# Pricing Using Subjectively Estimated Demand Curves

In situations when you don't know the price elasticity for a product or don't think you can rely on a linear or power demand curve, a good way to determine a product's demand curve is to identify the lowest price and highest price that seem reasonable. You can then try to estimate the product's demand with the high price, the low price, and a price midway between the high and low prices. This approach is based on a discussion in the book *Power Pricing*, by Robert Dolan. Given these three points on the product's demand curve, you can use the Microsoft Office Excel Trendline feature to fit a quadratic demand curve with the following equation:

$$(1)\ \text{Demand} = a(\text{price})^2 + b(\text{price}) + c$$

Fitting a quadratic demand curve in this manner enables the slope of the demand curve to either become steeper or flatter, which is much more realistic than the linear demand curve that requires the slope to remain constant.

For any three specified points on the demand curve, values of $a$, $b$, and $c$ exist that makes Equation 1 *exactly* fit the three specified points. Because Equation 1 fits three points on the demand curve, it seems reasonable to believe that the equation can give an accurate representation of demand for other prices. You can then use Equation 1 and Solver to maximize profit, which is given by the formula (price-unit cost)*demand. The following example shows how this process works.

Suppose that a drugstore pays $0.90 for each unit of ChapStick it orders. The store is considering charging from $1.50 through $2.50 for a unit of ChapStick. It thinks that at a price of $1.50, it can sell 60 units per week. (See the ChapStickprice.xls file.) At a price of $2.00, it thinks it can sell 51 units per week and at a price of $2.50, 20 units per week. To determine what price the store should charge for ChapStick, perform the following steps.

1.  Begin by entering the three points with which to chart your demand curve in the cell range E3:F6.
2.  Select E3:F6, click the Scatter option on the Charts group on the Ribbon, and then select the first option for a Scatter chart.

3. Right-click a data point and select Add Trendline.
4. In the Format Trendline dialog box, choose Polynomial, and select 2 in the Order box to obtain a quadratic curve of the form of Equation 1. Then select the option Display Equation on Chart. Figure 4-8 shows the required Trend Curve Settings. The chart containing the demand curve is shown in Figure 4-9.

   The estimated demand curve Equation 2 is as follows:

   $$(2)\ \text{Demand} = -44 * \text{Price}^2 + 136 * \text{Price} - 45$$

5. Next, insert a trial price in cell I2. Compute the product demand by using Equation 2 in cell I3 with the formula =-44*price^2+136*price-45. (Cell I2 is named Price.)
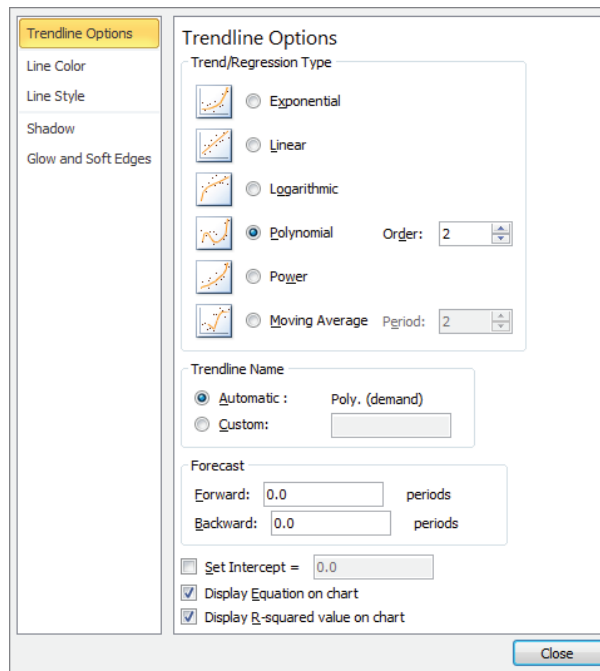


**Figure 4-8:** Pricing with a quadratic demand curve

6. Compute the weekly profit from ChapStick sales in cell I4 with the formula =demand*(price-unit_cost). (Cell E2 is named Unit_Cost and cell I3 is named Demand.)
7. Use Solver to determine the price that maximizes profit. The Solver Parameters dialog box is shown in Figure 4-10. The price is constrained to be from the lowest through the highest specified prices ($1.50 through $2.50). If you allow Solver to consider prices outside this range, the quadratic demand curve

might slope upward, which implies that a higher price would result in larger demand. This result is unreasonable, which is why you constrain the price.
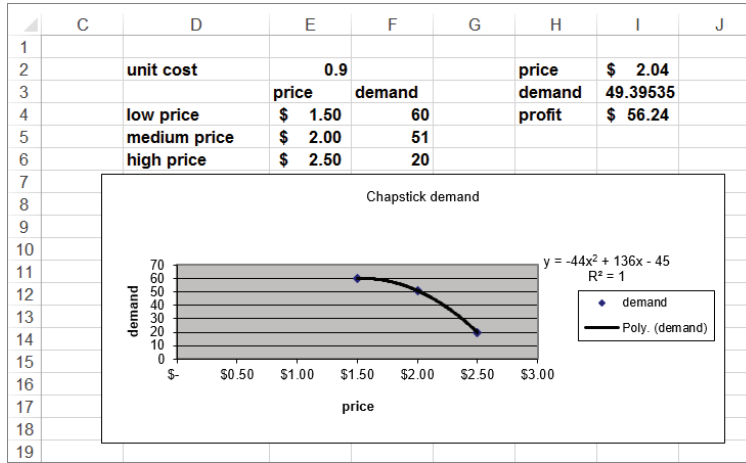


**Figure 4-9:** ChapStick demand curve
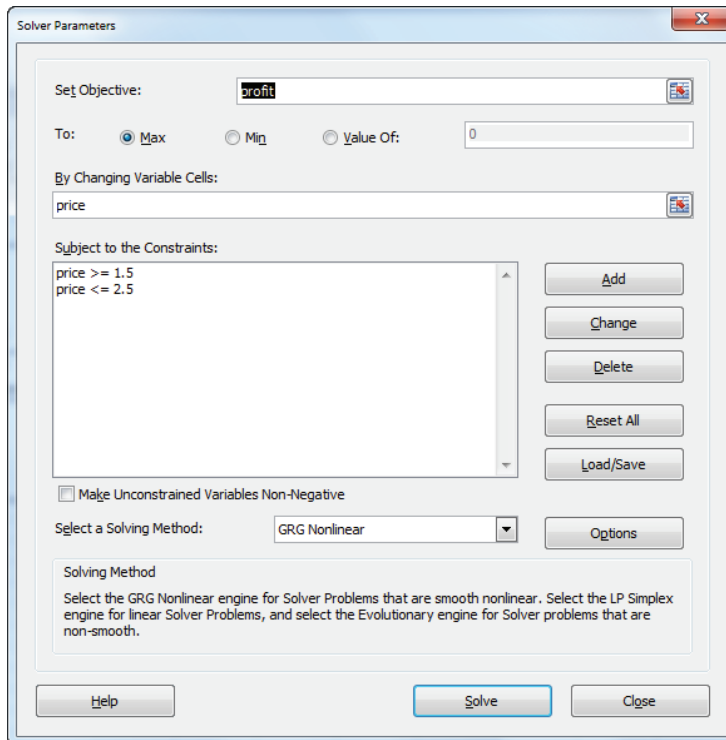


**Figure 4-10:** Solver window for quadratic demand curve example

You find that the drugstore should charge $2.04 for a ChapStick unit. This yields sales of 49.4 units per week and a weekly profit of $56.24.

**NOTE**    The approach to pricing outlined in this section requires no knowledge of the concept of price elasticity. Inherently, the Solver considers the elasticity for each price when it determines the profit-maximizing price.

**NOTE**    For the quadratic demand model to be useful, the minimum and maximum prices must be consistent with consumer preferences. A knowledgeable sales force should be able to come up with realistic minimum and maximum prices.

# Using SolverTable to Price Multiple Products

The approach developed to price a product in the preceding section can be extended to enable a company to easily price hundreds or thousands of products. The only information required for each product is the unit cost, estimated demand for lowest possible price, estimated demand for an intermediate price, and estimated demand. Using the SolverTable add-in written by Chris Albright (available for download at www.kelley.iu.edu/albright/Free_downloads.htm) you can easily fit the quadratic demand curve to each product. SolverTable is an Excel add-in that enables you to easily vary the inputs to a Solver model and track a wanted set of outputs.

After fitting the demand curves, you can use Solver to set a price for each product to maximize the total product generated from all products by completing the following steps (the Data worksheet in the Fittingmultipledemandcurves.xls file shows the estimated demand for three products at a low ($1.10), medium ($1.30), and high ($1.50) price):

1. Use an HLOOKUP function that keys off the entry in cell A11 to place the demands for each product in E14:E16. The Solver model chooses constant $a$, a coefficient $b$ for price, and a coefficient $c$ for price$^2$ that exactly passes through the demand points for each product.
2. Use SolverTable to loop through each product by changing the value in cell A11 in the range 1, 2, …, where $n$ = number of products.
3. Name the range F6:H10 Lookup.
4. Copy the formula =HLOOKUP($A$11,Lookup,C14) from E14 to E15:E16 to pull the demand for each product corresponding to the product index in A11.
5. Copy the formula =E$3+$E$4*D14+$E$5*(D14^2) from F14 to F15:F16 to compute the forecasted demand based on the quadratic demand curve based on the values in E3:E5.

6. Copy the formula =(E14-F14)^2 from G14 to G15:G16 to compute the squared error in the demand forecast for each price. For each price the estimation error is simply actual demand minus demand estimated from quadratic demand curve.

7. In cell G12 use the formula =SUM(G14:G16) to compute the sum of the squared estimation errors.

8. If you minimize G12 by changing E3:E5, Solver finds the values of *a*, *b*, and *c* that make the sum of squared errors equal to 0. You minimize the sum of squared errors instead of minimizing the sum of errors because if you do not square errors, then the sum of positive and sum of negative errors cancels each other out. Minimizing the sum of squared errors mimics the action of the Trend Curve Polynomial option described in the preceding section. The appropriate Solver Window is shown in Figure 4-11.
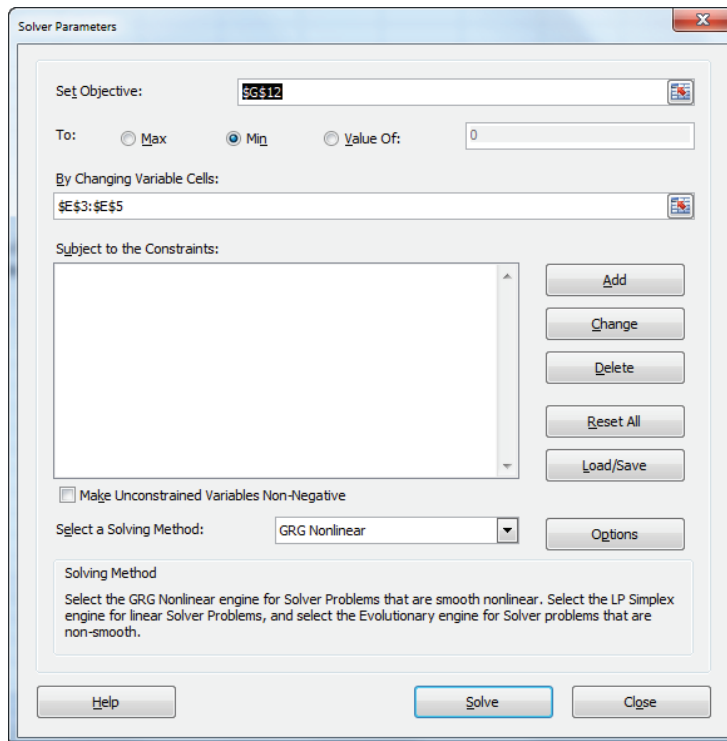


**Figure 4-11:** Solver window for SolverTable example

9. Choose the GRG Nonlinear option because the target cell is not constructed by adding up terms of the form (changing cell)*constant.

10. For Solver to obtain a correct solution to this problem, go to Options and check Use Automatic Scaling; then go to the GRG tab and select Central Derivatives.

11. Select Solve and you find the answer (see Figure 4-12). The Sum of Squared errors is 0 to 26 decimal places and the quadratic demand curve is $-73.625+195price+87.5price^2$. Because $a$, $b$, and $c$ can be negative, do not check the Make Unconstrained Variables Non-Negative box.

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | |
| 2 | | | | | | | | |
| 3 | | | | a | -73.625 | | | |
| 4 | | | | b | 195 | | | |
| 5 | | | | c | -87.5 | | | |
| 6 | | | | | | 1 | 2 | 3 |
| 7 | | | | | Price | Product 1 demand | Product 2 demand | Product 3 demand |
| 8 | | | | Low | 1.1 | 35 | 32 | 24 |
| 9 | | | | Medium | 1.3 | 32 | 27 | 17 |
| 10 | Data set | | | High | 1.5 | 22 | 16 | 9 |
| 11 | | 1 | | | | | SSE | |
| 12 | | | | | | | 1.00974E-27 | |
| 13 | | | | Price | Demand | Forecast | Squared Error | |
| 14 | | | 3 | 1.1 | 35 | 35 | 2.01948E-28 | |
| 15 | | | 4 | 1.3 | 32 | 32 | 8.07794E-28 | |
| 16 | | | 5 | 1.5 | 22 | 22 | 0 | |

**Figure 4-12:** Use of SolverTable for price optimization

## Using SolverTable to Find the Demand Curve for All Products

After you have a Solver model in a worksheet, you can use SolverTable to "loop through" the values of one (in a One-way SolverTable) or two (in a Two-way SolverTable) inputs and track any wanted outputs. To get started, assuming you have installed SolverTable, select SolverTable from the Ribbon, choose a One-way Table, and fill in the SolverTable dialog box, as shown in Figure 4-13.

To begin Solver place a 1 in the input cell (A11) and run the Solver to track the output cells ($a$, $b$, and $c$ and Sum of Squared Errors [SSE]). Then Solver places a 2 in A11 and finally a 3 in A11. The results are placed in a new worksheet (STS_1), as shown in Figure 4-14.
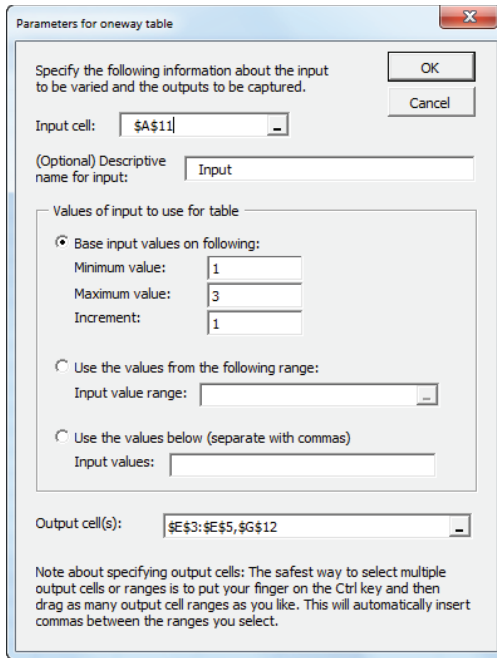
**Figure 4-13:** SolverTable settings

| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Oneway analysis for Solver model in data worksheet | | | | | | | | |
| 2 | | | | | | | | | 33.23375 |
| 3 | Input (cell $A$11) values along side, output cell(s) along top | | | | | | | | |
| 4 | | a | b | c | $G$12 | Price | demand | unit cost | Profit |
| 5 | 1 | -73.625 | 195 | -87.5 | 1.04441E-22 | 1.389493 | 28.3907 | 0.8 | 16.73611 |
| 6 | 2 | -47.75 | 155 | -75 | 6.31897E-25 | 1.370568 | 23.80379 | 0.9 | 11.2013 |
| 7 | 3 | 44.625 | -5 | -12.5 | 7.98128E-21 | 1.367619 | 14.40715 | 1 | 5.296336 |

**Figure 4-14:** SolverTable results

You can now set up a Solver model that determines the profit maximizing price for each product. Proceed as follows:

1. Enter trial prices for each product in the range F5:F7.
2. Based on the prices in F5:F7, compute the demand for each product by copying the formula =B5+C5*F5+F5^2*D5 from G5 to G6:G7.
3. Enter the unit cost for each product in H5:H7, and compute the profit for each product as *(price-unit cost)\*demand* by copying the formula =(F5-H5)*G5 from I5 to I6:I7.
4. Compute the total profit for all products in cell I2 with the formula =SUM(I5:I7).

5. Now use Solver (see Figure 4-15) to choose the prices in F5:F7 to maximize total profit (I2). This, of course, chooses the profit maximizing price for each product.
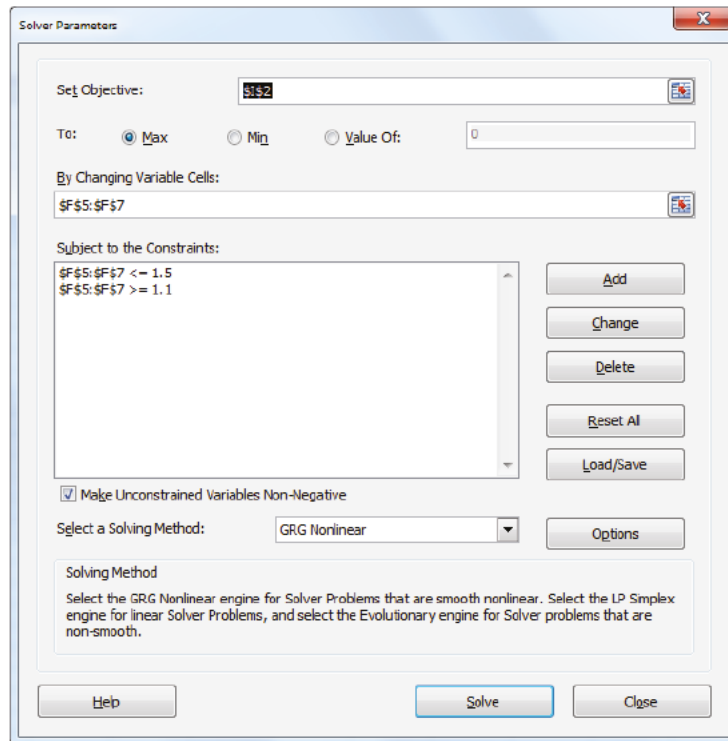


**Figure 4-15:** Finding profit maximizing price for each product

You find that you should charge $1.39 for Product 1 and $1.37 for Products 2 and 3. The SolverTable provides a powerful tool that can enable the marketing analyst to quickly estimate a demand curve for several products and determine a profit maximizing price for each product.

## Summary

In this chapter you learned the following:

■ Price elasticity of demand is the percentage decrease in demand for a 1 percent increase in price.

■ Given the current price and demand and price elasticity, you can use the Excel Trend Line feature to fit a linear demand curve (*Demand = a -b\*price*).

For a linear demand curve price, elasticity is different at every point on the demand curve.

■ Given the current price and demand and price elasticity, you can use the Excel Goal Seek feature to fit a power or constant elasticity demand curve ($Demand = a(price)^{-b}$)

■ Given a demand curve you can use Solver to find a profit maximizing price by maximizing $(price - unit\ cost) * demand$.

■ If you do not know the price elasticity, you can use the Polynomial portion of the Trend Line feature to fit a quadratic to three points on the demand curve.

■ Using the SolverTable add-in you can quickly fit the demand curve for a huge number of products and then use Solver to find the profit maximizing price for each product.

## Exercises

1. Your company charges $60 for a board game it invented and has sold 3,000 copies during the last year. Elasticity for board games is known to equal 3. Use this information to determine a linear and power demand curve.

2. You need to determine the profit-maximizing price for a video game console. Currently you charge $180 and sell 2 million consoles per year. It costs $150 to produce a console, and the price elasticity of demand for consoles is 3. What price should you charge for a console?

3. Now assume that, on average, a purchaser of your video game console buys 10 video games, and you earn $10 profit on each video game. What is the correct price for consoles?

4. In the `razorsandblades.xls` file example, suppose the cost to produce a blade is $0.20. If you charge $0.35 for a blade, a customer buys an average of 50 blades from you. Assume the price elasticity of demand for blades is 3. What price should you charge for a razor and for a blade?

5. You manage a movie theater that can handle up to 8,000 patrons per week. The current demand, price, and elasticity for ticket sales, popcorn, soda, and candy are given in Figure 4-16. The theater keeps 45 percent of ticket revenues. Unit cost per ticket, popcorn sales, candy sales, and soda sales are also given. Assuming linear demand curves, how can the theater maximize profits? Demand for foods is the fraction of patrons who purchase the given food.

**Figure 4-16:** Data for Exercise 5

6.  A prescription drug is produced in the United States and sold internationally. Each unit of the drug costs $60 to produce. In the German market, you sell the drug for 150 euros per unit. The current exchange rate is 0.667 U.S. dollars per euro. Current demand for the drug is 100 units, and the estimated elasticity is 2.5. Assuming a linear demand curve, determine the appropriate sales price (in euros) for the drug.

7.  Suppose it costs $250 to produce a video game console. A price between $200 and $400 is under consideration. Estimated demand for the game console is shown in the following table. What price should you charge for game console?

| Price | Demand |
| --- | --- |
| $200 | 50,000 |
| $300 | 25,000 |
| $400 | 12,000 |

8.  Use the demand information given in Exercise 7 for this exercise. Each game owner buys an average of 10 video games. You earn $10 profit per video game. What price should you charge for the game console?

9.  You want to determine the correct price for a new weekly magazine. The variable cost of printing and distributing a copy of the magazine is $0.50. You are thinking of charging from $0.50 through $1.30 per copy. The estimated weekly sales of the magazine are shown in the following table. What price should you charge for the magazine?

| Price | Demand (in Millions) |
| --- | --- |
| $0.50 | 2 |
| $0.90 | 1.2 |
| $1.30 | .3 |

10. Given the following information in the table for four products, find the profit-maximizing price for each product.

| Product | Lowest Price | Medium Price | High Price | Low Price Demand | Medium Price Demand | High Price Demand | Unit cost |
|---------|-----------|-------------|-----------|----------------|-------------------|-----------------|-----------|
| Product 1 | $1.40 | $2.20 | $3.00 | 100 | 40 | 5 | $1.10 |
| Product 2 | $2.20 | $3.00 | $4.00 | 200 | 130 | 25 | $1.50 |
| Product 3 | $45 | $75 | $95 | 400 | 300 | 130 | $40 |
| Product 4 | $200 | $250 | $300 | 600 | 300 | 50 | $120 |

11. (Requires calculus) Show that if the demand for a product is given by the power curve $q = ap^{-b}$, then for any price $a$, a 1 percent increase in price will decrease demand by 1 percent.

12. For the demand curve $q = 100p^{-2}$ show that for $p = 1, 2, 4, 8$, and $16$ that a 1 percent increase in price will result in approximately a 2 percent decrease in demand.

# 5

# Price Bundling

Chapter 4, "Estimating Demand Curves and Using Solver to Optimize Price," discussed situations in which a customer pays the same price for each unit she purchases of a product. When a customer pays the same price for each unit the seller is using *linear pricing*. This chapter and Chapter 6, "Nonlinear Pricing," describe some *nonlinear pricing* models in which the total amount a customer pays for a set of products is not equal to the sum of the individual product prices. One of the most common instances of nonlinear pricing is price bundling. This chapter shows that analytic models can help companies use bundling to maximize their profits.

## Why Bundle?

Companies often bundle products in an attempt to get customers to purchase more products than they would have without bundling. A few examples of bundling include the following:

- Cable companies bundle landlines, cell phone service, TV service, and Internet service.
- Automobile companies often bundle popular options such as navigation, satellite radio, and keyless entry.
- Computer mail order companies often bundle computers with printers, scanners, and monitors.
- Microsoft Office has been a highly successful bundling of software products such as Excel, Word, Access and Outlook.

In this chapter you will make the assumption that customers make decisions based on the concept of *consumer surplus*. The consumer surplus of a product is simply the value a consumer attaches to an available product minus the actual cost of the product.

> **NOTE** In Chapters 16, "Conjoint Analysis" and 18, "Discrete Choice Analysis," you learn how to determine the value a consumer attaches to a particular product or combination of products.

The consumer's goal is to make a choice that maximizes her consumer surplus. Therefore, if each product combination has a negative surplus, no product combination is purchased; however, as long as at least one option has a non-negative consumer surplus, the consumer will choose the option with the largest consumer surplus. The possibility of ties can be ignored. This section examines three examples that portray how bundling can increase profits.

## Bundling Products to Extract Consumer Surplus

Suppose you own Disney World and each customer values the five rides in the park as follows:

- Space Mountain: $10
- Matterhorn: $8
- Tower of Terror: $6
- Small World: $4
- Mr. Toad: $2

How can Disney World maximize revenue from these rides? First assume a single price is charged for each ride. If Disney World charges $10 a ride, each customer would go on only Space Mountain and you make $10 per customer. If Disney World charges $8 per ride, each customer would go on two rides, and you make $16. If you charge $6 per ride, then each customer goes on three rides and you make $18. If you charge $4 per ride, then each customer goes on four rides and you make $16. Finally, charging $2 per ride, you make $10 in revenue. Therefore with a single price for each ride, revenue is maximized at $18 per customer.

Now suppose Disney World does not offer per ride tickets but simply offers a five-ride ticket for $30. Because the consumer values all five rides at $30 (the sum of the ride values) he will pay $30 for park admission. This is a 67 percent increase in revenue. Of course, single park admission is the strategy Disney World has adapted. Single park admission has other benefits such as reducing lines, but this simple example shows that product bundling can help companies "extract" consumer surplus from customers. Later in the chapter you see that the assumption that all customers are identical is not important to your analysis.

## Pure Bundling

Now consider another example. Suppose you sell two products and have two potential customers that value each product, as shown in Table 5-1:

**Table 5-1:** Pure Bundling Works Well with Negatively Correlated Valuations

| Product | Customer 1 | Customer 2 |
|---------|------------|------------|
| Computer | $1,000 | $500 |
| Monitor | $500 | $1,000 |

For simplicity, assume the cost of producing each product is negligible and can be assumed to equal $0. Therefore revenue and profit are identical. In this example customer valuations are negatively correlated; Customer 1 values the computer more than Customer 2 but values the monitor less. If you price each product separately, you would charge $1,000 for each product and make $2,000 revenue; however, if you charge $1,500 for a bundle of both products, you make $3,000. Thus, when customer valuations for products are negatively correlated, bundling can result in a significant increase in profit.

The reason bundling works in this case is that bundling enables the seller to entirely extract the value the consumers attach to the products. If the seller only offers the customer a choice between purchasing all products or nothing, the situation is called *pure bundling*. Movie rental companies usually give theatres a choice between renting an assortment of some blockbuster movies and some not-so-popular movies, or renting no movies at all. In this case if the theatre owner wants the blockbuster movies she has to also rent the less-popular movies.

If customer valuations are positively correlated then bundling usually offers no benefit, because bundling does not allow the seller to extract more surplus than separate prices. For example, in Table 5-2 valuations are positively correlated because Customer 2 is willing to pay more for each product than Customer 1. In this case, if the seller offers only separate prices she would charge $1,000 for the computer and $500 for the monitor. Then each customer buys both products and total revenue is $3,000. If the seller offers a pure bundle she should charge $1,500 for the bundle. Then both customers purchase the bundle for $1,500 yielding a total revenue of $3,000. In this this case pure bundling and separate prices yield identical profits.

**Table 5-2:** Pure Bundling Works Poorly with Positively Correlated Valuations

| Product | Customer 1 | Customer 2 |
|---|---|---|
| Computer | $1,000 | $1,250 |
| Monitor | $500 | $750 |

## Mixed Bundling

*Mixed bundling* means the seller offers a different price for each available combination of products. For the data in Table 5-3 mixed bundling is optimal.

**Table 5-3:** Mixed Bundling is Optimal

| Customer | Computer | Monitor |
|---|---|---|
| 1 | $900 | $800 |
| 2 | $1,100 | $600 |
| 3 | $1,300 | $400 |
| 4 | $1,500 | $200 |
| Unit Cost | $1,000 | $300 |

To maximize profits, the seller can do the following (see also Exercise 7):

- To maximize profit from separate prices the seller charges $600 for the monitor and $1,300 for the computer. In this scenario Customers 3 and 4 buy the computer and Customers 1 and 2 buy the monitor yielding a profit of $1,200.
- To maximize profit with mixed bundling the seller can charge $799 for the monitor, $1,499 for the computer, and $1,700 for the bundle. Then Customer 1 buys just a monitor, Customer 4 buys just a computer, and Customers 2 and 3 buy the product bundle. The seller earns a profit of $1,798.
- If only the monitor and computer pure bundle is offered, the seller should sell the pure bundle for $1,700. Then each customer will purchase the pure bundle and the seller earns a profit of $1,600.

# Using Evolutionary Solver to Find Optimal Bundle Prices

When you combine the consumer surplus decision-making assumptions explained in the previous section with the power of the Evolutionary Solver Engine of Excel 2010 or later, you can easily solve complex bundling problems.

For this section's example, suppose Verizon sells cell phone service, Internet access, and FIOS TV service to customers. Customers can buy any combination of these three products (or not buy any). The seven available product combinations include the following:

- Internet (Combination 1)
- TV (Combination 2)
- Cell phone (Combination 3)
- Internet and TV (Combination 4)
- Internet and cell phone (Combination 5)
- TV and cell phone (Combination 6)
- All three products (Combination 7)

The file phone.xls gives the amount 77 representative customers are willing to pay per month for each service. Use the model in the initial solver worksheet to follow along with the example (also shown in Figure 5-1). In row 6 you see that the first customer is willing to pay up to $3.50 per month for Internet, $7 per month for TV service, and $3.50 per month for cell phone service. (It's hard to make money off this customer!) You can use the Evolutionary Solver Engine and the willingness to pay data to determine a price for each product combination that maximizes revenue.

| | Internet | TV | Cell Phon | Internet | TV | Cell | Internet +TV | Internet +Cell | TV+Cel I | All 3 | max surp | bought? | revenue | | Product | Frequenc |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Product | | | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | total | 3413.696 | | | | |
| Price | | | | 74.35303 | 35 | 82.16 | 69.9942 | 69.99101 | 69.991 | 89.945 | | | | | | |
| | $3.50 | $7.00 | $3.50 | -70.853 | -28 | -78.7 | -59.494 | -62.991 | -59.491 | -75.945 | -27.9987 | 0 | 0 | | | |
| | $17.50 | $35.00 | $3.50 | -56.853 | 0 | -78.7 | -17.494 | -48.991 | -31.491 | -33.945 | 0.0013 | 2 | 34.9987 | | | |
| | $28.00 | $28.00 | $49.00 | -46.353 | -7 | -33.2 | -13.994 | 7.008987 | 7.0087 | 15.055 | 15.05482 | 7 | 89.94518 | | Product | Frequenc |
| | $70.00 | $70.00 | $0.00 | -4.35303 | 35 | -82.2 | 70.0058 | 0.008987 | 0.0087 | 50.055 | 70.00576 | 4 | 69.99424 | | 0 | 25 |
| | $0.00 | $7.00 | $14.00 | -74.353 | -28 | -68.2 | -62.994 | -55.991 | -48.991 | -68.945 | -27.9987 | 0 | 0 | | 1 | 0 |
| | $0.00 | $70.00 | $0.00 | -74.353 | 35 | -82.2 | 0.00576 | -69.991 | 0.0087 | -19.945 | 35.0013 | 2 | 34.9987 | | 2 | 19 |
| | $21.00 | $35.00 | $10.50 | -53.353 | 0 | -71.7 | -13.994 | -38.491 | -24.491 | -23.445 | 0.0013 | 2 | 34.9987 | | 3 | 0 |
| | $7.00 | $21.00 | $0.00 | -67.353 | -14 | -82.2 | -41.994 | -62.991 | -48.991 | -61.945 | -13.9987 | 0 | 0 | | 4 | 8 |
| | $5.25 | $7.00 | $2.10 | -69.103 | -28 | -80.1 | -57.744 | -62.641 | -60.891 | -75.595 | -27.9987 | 0 | 0 | | 5 | 2 |
| | $21.00 | $28.00 | $28.00 | -53.353 | -7 | -54.2 | -20.994 | -20.991 | -13.991 | -12.945 | -6.9987 | 0 | 0 | | 6 | 1 |
| | $35.00 | $49.00 | $21.00 | -39.353 | 14 | -61.2 | 14.0058 | -13.991 | 0.0087 | 15.055 | 15.05482 | 7 | 89.94518 | | 7 | 22 |
| | $21.00 | $21.00 | $21.00 | -53.353 | -14 | -61.2 | -27.994 | -27.991 | -27.991 | -26.945 | -13.9987 | 0 | 0 | | | |
| | $14.00 | $35.00 | $21.00 | -60.353 | 0 | -61.2 | -20.994 | -34.991 | -13.991 | -19.945 | 0.0013 | 2 | 34.9987 | | | |
| | $14.00 | $35.00 | $14.00 | -60.353 | 0 | -68.2 | -20.994 | -41.991 | -20.991 | -26.945 | 0.0013 | 2 | 34.9987 | | | |
| | $70.00 | $0.00 | $49.00 | -4.35303 | -35 | -33.2 | 0.00576 | 49.00899 | -20.991 | 29.055 | 49.00899 | 5 | 69.99101 | | | |
| | $7.00 | $35.00 | $14.00 | -67.353 | 0 | -68.2 | -27.994 | -48.991 | -20.991 | -33.945 | 0.0013 | 2 | 34.9987 | | | |
| | $21.00 | $35.00 | $42.00 | -53.353 | 0 | -40.2 | -13.994 | -6.99101 | 7.0087 | 8.0548 | 8.054824 | 7 | 89.94518 | | | |

**Figure 5-1:** Verizon Internet example

The key to your model is to set up a spreadsheet that tells you, for any set of prices for each possible product combination, how much revenue you can obtain from your sample of customers. Then you can use the Evolutionary Solver to find the set of prices for the product combinations that maximize your revenue. To find how much revenue you can generate for any set of product combination prices, proceed as follows:

1. In D4:J4 enter trial prices for each of the possible seven product combinations.

2. In Row 6, determine the first's customer's consumer surplus by computing her value for the products in a combination and subtracting the cost of the product combination. For example, the first customer's consumer surplus for the Internet +TV product combination is computed in cell G6 with the formula =A6+B6-G$4. Copy the formulas in row 6 to the cell range D7:J82 to compute each customer's consumer surplus for each product combination.

3. Determine for the set of prices in row 4 which product combination, if any, is purchased. Copy the formula =MAX(D6:J6) from K6 to K7:K82 to find each consumer's surplus.

4. Now here's the key to your spreadsheet. In Column L you use the MATCH function (introduced in Chapter 2, "Using Excel Charts to Summarize Marketing Data") combined with an IF statement to determine which product combination each customer purchases. Use product combination 0 to denote no purchase, whereas the actual product combinations are indicated by the integers 1–7. Copy the formula =IF(K6<0,0,MATCH(K6,D6:J6,0)) from L6 to L7:L82 yields the product combination (if any) bought by each customer.

5. Copy the formula =IF(L6=0,0,HLOOKUP(L6,$D$3:$J$4,2)) from M6 to M7:M82 to compute for each person the revenue generated.

6. In cell M4, compute your total revenue with the formula =SUM(M6:M82).

To find the product combination prices that maximize revenue you need to use the Evolutionary Solver, so some discussion of the Evolutionary Solver is in order.

## Introduction to the Evolutionary Solver

As explained in Chapter 4, "Estimating Demand Curves and Using Solver to Optimize Price," the Evolutionary Engine is used when your target cell and constraints contain nonsmooth functions that reference changing cells. For example, if your target cell and constraints contain IF, SUMIF, COUNTIF, SUMIFS, COUNTIFS, AVERAGEIF, AVERAGEIFS, ABS, MAX, or MIN functions that reference the changing cells,

then the Evolutionary engine probably has the best chance to find a good solution to your optimization problem. The model makes extensive use of IF statements, so it is a good choice to use the Evolutionary Solver. The target cell is to maximize revenue (cell M3) by changing the product prices (cell range D4:J4).

When using the Evolutionary Solver, you should follow these rules:

- Place upper and lower bounds on each of your changing cells. This makes the Solver's search for an optimal solution much easier. In the Verizon bundling model use a lower bound of 0 for each price and an upper bound of 100.

**NOTE** If, when running Evolutionary Solver, a changing cell hits an upper or lower bound, you should relax that bound because Solver was trying to push the changing cell outside the range you specified.

- Use no constraints other than the bounds on the changing cells. You will soon see how to use penalties in the target cell to enforce constraints.
- In the Options tab increase Max Time to 3600 seconds. This causes Solver to run for 60 minutes, even if you leave the room. This should be plenty of time for Solver to find a good solution. Also add a few 000s to Max Subproblems and Max Feasible Solutions. This ensures that Solver does not stop when you leave the room. Figure 5-2 summarizes these settings.
- In the Options tab select the Evolutionary Solver tab and change Mutation Rate to 0.5 and Maximum Time Without Improvement to 300 seconds. (Mutation rate is explained soon.) Setting a Maximum Time Without Improvement to 5 minutes ensures that if Solver fails in 5 minutes to improve the current solution, the Solver stops. Hitting the Escape key at any time stops the Solver.

*Evolutionary algorithms* (often called *genetic algorithms*) were discovered by John Holland, a Michigan computer science processor. Evolutionary Solver begins with between 50 and 200 "solutions," which are randomly chosen values of the changing cells that lie between the lower and upper bounds for each changing cell. The exact number of solutions used is specified in the Population Size field in Figure 5-3. The default value is 100, which is fine. Then the target cell value is computed for each solution. By a process of "reproduction" explained in David Goldberg's textbook *Genetic Algorithms* (Addison-Wesley, 1989), a new set of 100 solutions is obtained. Previous solutions that have "good" values (large in a Max problem and small in a Min problem) have a better chance of surviving to the next generation of solutions. This is the mathematical implementation of Darwin's survival of the fittest principle.
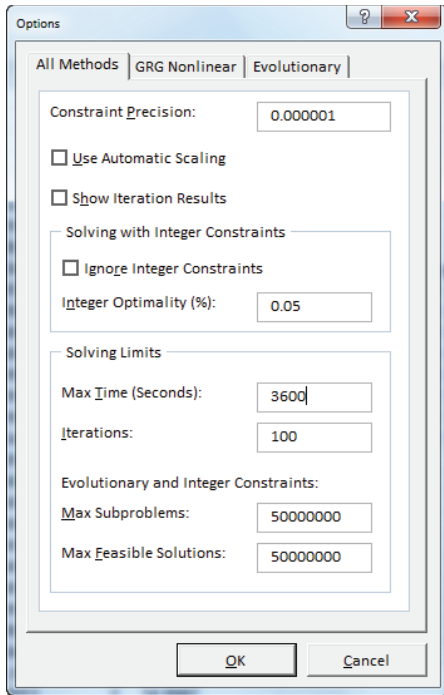
**Figure 5-2:** Evolutionary Solver settings

The Excel Solver also enables solutions to improve via Mutation. To understand how Mutation works you need to know that Excel codes each value of a changing cell in binary notation. For example, the number 9 is coded as 1001 ($1 * 2^3 + 0 * 2^2 + 0 * 2^1 + 1$).

A Mutation in the Evolutionary Solver changes a 0 to a 1 or a 1 to a 0. A higher value for the Mutation rate moves you around more in the feasible region but exposes you to a larger probability of being led on a wild goose chase after moving to a part of the feasible region with poor changing cell values. A Mutation rate of .5 usually yields the best results. Figure 5-3 shows the Mutation rate changed to .5. The amazingly simple Evolutionary Solver search procedure based on survival of the fittest can solve extraordinarily complex problems! When solving a problem the Evolutionary Solver must make many random choices. This means that if two different people run the same model, then after 5 minutes or so they may see different optimal solutions. Eventually they should see target cell values that are virtually identical.
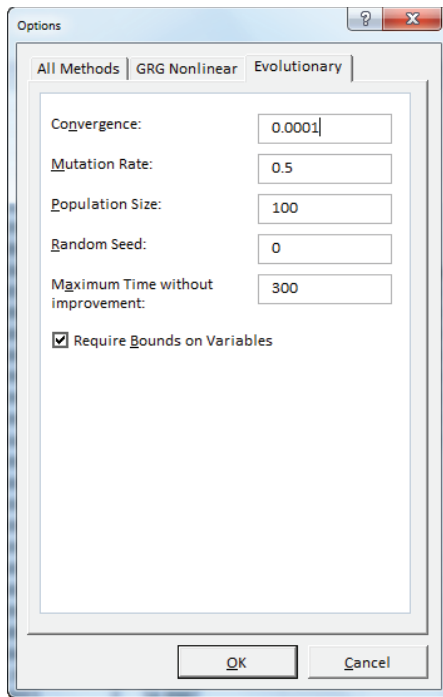
**Figure 5-3:** Changing the Mutation rate

## Finding the Optimal Bundle Prices

The Solver window shown in Figure 5-4 enables you to find revenue maximizing the set of prices for each product combination. A maximum profit of $3,413.70 (see worksheet `initial solver`) is found with the product combination prices, as shown in Figure 5-5.

**NOTE**    Unlike the GRG or Simplex Solver engine, the Evolutionary Solver engine is only guaranteed to find a near optimal solution (as opposed to an optimal solution). Therefore, when you run the Evolutionary Solver on the book's examples, do not be surprised if your optimal target cell value differs slightly from the target cell value shown in the book.
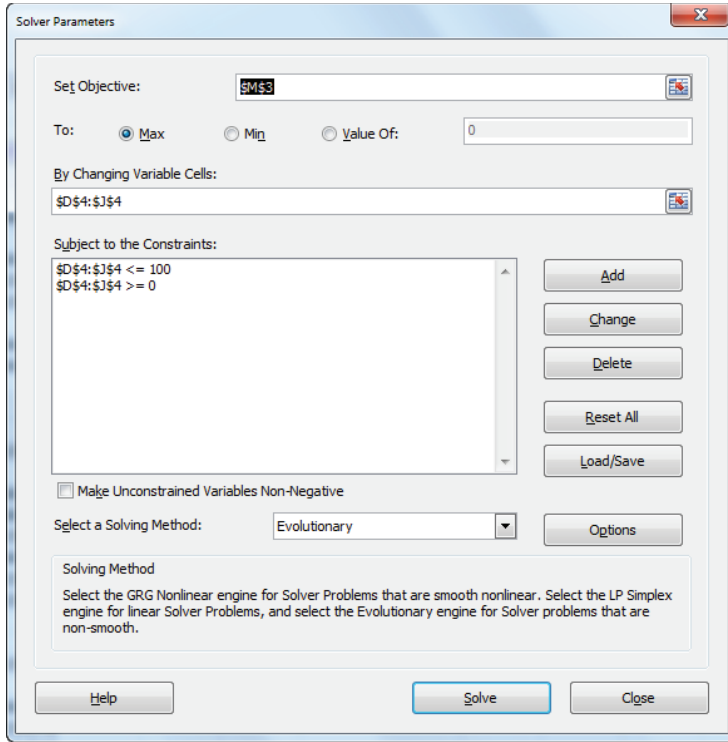
**Figure 5-4:** Initial bundling Solver model

| C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|
| 3 Product | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 4 Price | 74.35303 | 35 | 82.16 | 69.9942 | 69.99101 | 69.991267 | 89.945 |
| 5 Cell Phone | Internet | TV | Cell | Internet +TV | Internet +Cell | TV+Cell | All 3 |

**Figure 5-5:** Solution to initial bundling model

Unfortunately, you could not go to market charging $74.35 for Internet service and $69.99 for Internet + TV service because it is unreasonable to provide two services for a lower price than a single service. This situation is called a *price reversal*. As you soon see, the reason Solver charged a higher price for Internet service than for Internet and TV service is because the revenue maximizing prices involve nobody just buying Internet service, so a high price helps ensure that nobody buys just Internet service.

Take a look at the `final solver` worksheet, and you see it contains the same prices found in the `initial solver` worksheet. The cell range P70:Q81 in the `final`

`solver` worksheet (as shown in Figure 5-6) computes the price reversals for each possible combination of a product with a larger combination of products. In column P, the price of the smaller product combination is subtracted from the price of the larger product combination. For example, in P71 the formula =D4-G4 computes the Internet Price – (Internet + TV Price). The difference of $4.36 indicates that the Internet + TV price is $4.36 lower than the Internet price. To ensure that no price reversals occur, the target cell is penalized $500 for each dollar of price reversal. Then, survival of the fittest should ensure that no price reversals survive. The choice of penalty amount is an art, not a science. Too large a penalty (such as $1 million) seems to work poorly, and too small a penalty does not "kill off" the things you want to avoid.

| | O | P | Q |
|---|---|---|---|
| 68 | | | |
| 69 | Penalty | dev | penalty |
| 70 | TV-(I+TV) | -34.9955 | 0 |
| 71 | I-(I+TV) | 4.358796 | 4.358796 |
| 72 | Cell-(I+Cell) | 12.16776 | 12.16776 |
| 73 | I-(I+C) | 4.36202 | 4.36202 |
| 74 | TV-(TV+Cell) | -34.9926 | 0 |
| 75 | Cell-(TV+Cell) | 12.1675 | 12.1675 |
| 76 | (I+TV)-All | -19.9509 | 0 |
| 77 | (I+C)-All | -19.9542 | 0 |
| 78 | (TV+Cell)-All | -19.9539 | 0 |
| 79 | TV-All | -54.9465 | 0 |
| 80 | Cell-All | -7.7864 | 0 |
| 81 | I-All | -15.5921 | 0 |
| 82 | Total | | 33.05608 |

**Figure 5-6:** Price reversals for initial Solver solution

You can copy the formula =IF(P70>0,P70,0) from Q70 to Q71:Q82, to track the price reversal because it yields a value of 0 if the product comparison does not have a price reversal. With the formula =SUM(M6:M82)-500*Q82 you can incorporate your penalty in the target cell. The Solver window is the same (refer to Figure 5-4). After running the Solver you can find the optimal solution, as shown in Figure 5-7.

Therefore, after completing all the calculations, a maximum profit of $3,413.90 is obtained by charging the following:

- $47.41 for Internet service
- $35 for TV service
- $67.87 for cell phone service
- $70 for any two product combination
- $89.95 for all three products

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | total | 3413.9 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Price** | 47.41 | 35 | 67.87 | 70 | 70 | 70 | 89.95 | | | |
| **Cell Phon**/Intern et | | TV | Cell | Internet +TV | Internet +Cell | TV+Cel l | All 3 | max surp | bought? | revenue |
| $3.50 | -43.91 | -28 | -64.4 | -59.5 | -63 | -59.5 | -75.95 | -28 | 0 | 0 |
| $3.50 | -29.91 | 0 | -64.4 | -17.5 | -49 | -31.5 | -33.95 | 1.39E-09 | 2 | 35 |
| $49.00 | -19.41 | -7 | -18.9 | -14 | 7.000001 | 7 | 15.05 | 15.05 | 7 | 89.95 |
| $0.00 | 22.59 | 35 | -67.9 | 70 | 6.01E-07 | 5E-08 | 50.05 | 70 | 4 | 70 |
| $14.00 | -47.41 | -28 | -53.9 | -63 | -56 | -49 | -68.95 | -28 | 0 | 0 |
| $0.00 | -47.41 | 35 | -67.9 | 1.9E-08 | -70 | 5E-08 | -19.95 | 35 | 2 | 35 |
| $10.50 | -26.41 | 0 | -57.4 | -14 | -38.5 | -24.5 | -23.45 | 1.39E-09 | 2 | 35 |
| $0.00 | -40.41 | -14 | -67.9 | -42 | -63 | -49 | -61.95 | -14 | 0 | 0 |
| $2.10 | -42.16 | -28 | -65.8 | -57.75 | -62.65 | -60.9 | -75.6 | -28 | 0 | 0 |
| $28.00 | -26.41 | -7 | -39.9 | -21 | -21 | -14 | -12.95 | -7 | 0 | 0 |
| $21.00 | -12.41 | 14 | -46.9 | 14 | -14 | 5E-08 | 15.05 | 15.05 | 7 | 89.95 |
| $21.00 | -26.41 | -14 | -46.9 | -28 | -28 | -28 | -26.95 | -14 | 0 | 0 |

**Figure 5-7:** Final bundling solution

The consumer is given a substantial discount if she buys two or more products. You can copy the formula =COUNTIF($L$6:$L$82,N9) from O9 to O10:O16 to find the number of people purchasing each product combination. As shown in Figure 5-8, 25 people buy nothing; nobody buys just the Internet; 19 buy just TV; nobody buys just the cell phone; 8 buy the Internet + TV; 2 buy Internet + cell phone; 1 buys TV + cell phone; and 22 people buy the bundle! Your pricing helped extract the high value people place on the Internet and cell phone service by incentivizing these people to buy more than one service. Note the high prices for the Internet and cell phone are designed to give customers an incentive to purchase more products.

| Product | Frequenc |
|---|---|
| 0 | 25 |
| 1 | 0 |
| 2 | 19 |
| 3 | 0 |
| 4 | 8 |
| 5 | 2 |
| 6 | 1 |
| 7 | 22 |

**Figure 5-8:** Number of purchases of each product combination

This method will automatically determine whether separate prices, mixed bundling, or pure bundling is optimal.

## Summary

In this chapter you learned the following:

- Bundling products often allows companies to extract more consumer surplus from customers by incentivizing the customers to purchase more products.
- If you assume that each customer will purchase the product with the largest (non-negative) surplus, then you can set up a spreadsheet that tells you the revenue obtained from your customers for any set of prices.
- You can use Evolutionary Solver to maximize Revenue (or Profit) obtained from customers.
- When using Evolutionary Solver you need to implement bounds on changing cells, use a Mutation rate of 0.5, and handle other constraints by incorporating penalties in the target cell.

## Exercises

1. A German machine company sells industrial machinery and maintenance policies for the machine. There are four market segments. Figure 5-9 shows the size of each market segment and the amount each segment will pay for the machine separately, maintenance separately, or the bundle of machinery + maintenance. The variable cost is $550 per machine and $470 per maintenance agreement. What set of prices maximize profits?

|   | B | C | D | E |
|---|---|---|---|---|
| 1 | **fixed** | Machine | Maintenance | Bundle |
| 2 | Costs | 550 | 470 | 1020 |
| 3 | Size | Machine Use Max | Maintenance Max | Bundle Max |
| 4 | 12 | 1250 | 1500 | 2600 |
| 5 | 23 | 1450 | 540 | 1750 |
| 6 | 22 | 1080 | 1200 | 2090 |
| 7 | 43 | 1390 | 1100 | 2350 |

**Figure 5-9** Data for Exercise 1

2. The file songdata.xlsx gives the values several hundred people place on 15 downloadable songs. A subset of the data is shown in Figure 5-10.

a. If you charge a single price for all songs, how can you maximize revenue?
b. If you can charge two different prices for songs, how can you maximize revenue?

| | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|
| 10 | Person | Song 1 | Song 2 | Song 3 | Song 4 | Song 5 | Song 6 | Song 7 |
| 11 | 1 | $0.76 | $0.68 | $1.61 | $1.82 | $0.26 | $0.66 | $1.51 |
| 12 | 2 | $0.74 | $0.89 | $1.78 | $1.87 | $0.49 | $0.68 | $1.42 |
| 13 | 3 | $0.36 | $0.49 | $1.40 | $1.42 | -$0.02 | $0.33 | $1.14 |
| 14 | 4 | $0.34 | $0.45 | $1.42 | $1.50 | $0.07 | $0.34 | $1.14 |
| 15 | 5 | $1.27 | $1.31 | $2.14 | $2.38 | $1.02 | $1.19 | $2.05 |
| 16 | 6 | $1.46 | $1.55 | $2.53 | $2.52 | $1.15 | $1.38 | $2.18 |
| 17 | 7 | $0.57 | $0.70 | $1.51 | $1.48 | $0.11 | $0.46 | $1.23 |
| 18 | 8 | $0.62 | $0.74 | $1.48 | $1.71 | $0.31 | $0.47 | $1.29 |
| 19 | 9 | $1.37 | $1.50 | $2.31 | $2.57 | $1.09 | $1.31 | $2.14 |
| 20 | 10 | $0.47 | $0.38 | $1.39 | $1.51 | $0.10 | $0.31 | $1.11 |
| 21 | 11 | $1.50 | $1.61 | $2.42 | $2.56 | $1.13 | $1.43 | $2.10 |
| 22 | 12 | $0.57 | $0.73 | $1.53 | $1.65 | $0.37 | $0.62 | $1.36 |
| 23 | 13 | $0.92 | $0.87 | $1.76 | $1.82 | $0.51 | $0.82 | $1.63 |
| 24 | 14 | $1.02 | $1.13 | $2.07 | $2.22 | $0.65 | $1.11 | $1.91 |
| 25 | 15 | $1.27 | $1.32 | $2.10 | $2.28 | $0.77 | $1.23 | $1.94 |

**Figure 5-10** Data for Exercise 2

3. Kroger is trying to determine which types of Vanilla Wafer cookies to stock. The wholesale cost Kroger is charged for a box of each type of cookie is shown in the following table:

| National 1 | National 2 | Generic | Copycat |
|---|---|---|---|
| $1.5 | $2.00 | $0.90 | $1.25 |

The wholesale price of a box of cookies is assumed equal to the product's quality. Each customer assigns a different value to quality, as shown in file Wafersdata.xlsx and Figure 5-11. For example, Customer 7 would value National 1 at 1.02 * (1.5), and so on. Your task is to determine how to price each type of cookie and then to recommend which brands Kroger should stock to maximize profit from cookies.

| | E | F | G |
|---|---|---|---|
| 7 | Customer | Quality value | |
| 8 | 1 | 1 | |
| 9 | 2 | 1.003333 | |
| 10 | 3 | 1.006667 | |
| 11 | 4 | 1.01 | |
| 12 | 5 | 1.013333 | |
| 13 | 6 | 1.016667 | |
| 14 | 7 | 1.02 | |
| 15 | 8 | 1.023333 | |
| 16 | 9 | 1.026667 | |
| 17 | 10 | 1.03 | |
| 18 | 11 | 1.033333 | |
| 19 | 12 | 1.036667 | |
| 20 | 13 | 1.04 | |
| 21 | 14 | 1.043333 | |
| 22 | 15 | 1.046667 | |
| 23 | 16 | 1.05 | |

**Figure 5-11** Valuations for Exercise 3

4.  Microsoft is going to sell a Student version of Office (excluding Access and Outlook) and a full version of Office. Assume there are three market segments whose size and valuations for each version of Office are given in the following table. What price for each version of Office can maximize Microsoft's revenue?

| Segment | Student Version | Full Version | Size |
|---|---|---|---|
| Students | 110 | 160 | 20% |
| Individuals (nonstudent) | 170 | 300 | 40% |
| Businesses | 240 | 500 | 40% |

5.  *The New York Times* wants to price home subscriptions and web access. Of course, people can buy both home subscriptions and web access. Assume there are three market segments whose size and valuations (per month) for each product combination are given here. The cost to provide a home subscription is $15 and the variable cost to provide web access is $1.

| Segment | Home Subscription | Web Access | Home +Web | Size |
|---|---|---|---|---|
| 1 | $30 | $35 | $50 | 25% |
| 2 | $35 | $15 | $40 | 40% |
| 3 | $20 | $20 | $25 | 35% |

What prices maximize monthly profits?

6. Before publishing a hardcover book, a publisher wants to determine the proper price to charge for the hardcover and paperback versions. The file `paperback` `.xlsx` contains the valuations of 50 representative potential purchasers of the book. Suppose the unit cost of the hardcover book is $4 and the unit cost of the paperback is $2. If the bookstore charges double what it pays the publisher, what price should the publisher charge the bookstore for each version of the book?

7. For the data in Table 5-3 verify that the given mixed bundling strategy maximizes the seller's profit.

8. Describe a situation in which a bank can benefit from a bundling strategy.

9. Suppose your company sells four products and during a year a customer will buy each product at most once. Your company would like to give an end-of-year rebate of 10 percent off to customers who purchase at least $1,000 of products. How can the methods of this chapter be used to maximize your company's profits?

# 6

# Nonlinear Pricing

**N**onlinear pricing is used in a situation in which the cost of purchasing $q$ units is not equal to the unit cost $c$ per unit times $q$. Bundling (discussed in Chapter 5, "Price Bundling") is a special type of nonlinear pricing because the price of three items is not equal to the sum of the individual prices.

Other common examples of nonlinear pricing strategies include:

- **Quantity discounts:** If customers buy ≤$CUT$ units, they pay high price (*HP*) per unit, and if they buy >$CUT$ units, they pay low price (*LP*) per unit. *CUT* is simply the "cutoff point at which the charged price changes." For example, if customers buy ≤*1000* units, you charge $10 a unit, but if they buy more than 1000 units, you charge $8 per unit for all units bought. This form of nonlinear pricing is called the *nonstandard quantity discount*. Another type of quantity discount strategy is as follows: Charge *HP* for the first *CUT* units bought, and charge *LP* for remaining units bought. For example, you charge $10 per unit for first 1000 units and $8 per unit for remaining units bought. This form of nonlinear pricing is called the *standard quantity discount*. In both examples the value of *CUT = 1000*.

- **Two-part tariff:** The cost of buying $q$ units is a fixed charge $K$ plus $$c$ per unit purchased. For example, it may cost $500 to join a golf club and $30 per round of golf.

Many companies use quantity discounts and two-part tariffs. Microsoft does not charge twice as much for 200 units, for example, as for 100 units. Supermarkets charge less per ounce for a 2-pound jar of peanut butter as for a 1-pound jar of peanut butter. Golf courses often use a two-part tariff by charging an annual membership fee and a charge for each round of golf.

Just as in Chapter 5 you used the Evolutionary Solver to find optimal price bundling strategies, you can use the Evolutionary Solver to find the profit maximizing parameters of a nonlinear pricing strategy. As in Chapter 5, you can assume the

consumer will choose an option giving her the maximum (non-negative) consumer surplus. You can see that nonlinear pricing can often significantly increase your profits, seemingly creating profits out of nothing at all.

In this chapter you will first learn how a consumer's demand curve yields the consumer's willingness to pay for each unit of a product. Using this information you will learn how to use the Evolutionary Solver to determine profit or revenue maximizing nonlinear pricing strategies.

## Demand Curves and Willingness to Pay

A demand curve tells you for each possible price how many units a customer is willing to buy. A consumer's willingness to pay curve is defined as the maximum amount a customer is willing to pay for each unit of the product. In this section you will learn how to extract the willingness to pay curve from a demand curve.

Suppose you want to sell a software program to a Fortune 500 company. Let $q$ equal the number of copies of the program the company demands, and let $p$ equal the price charged for the software. Suppose you estimated that the demand curve for software is given by $q = 400 - p$. Clearly, your customer is willing to pay less for each additional unit of your software program. Locked inside this demand curve is information about how much the company is willing to pay for each unit of your program. This information is crucial to maximize profitability of sales.

Now rewrite your demand curve as $p = 400 - q$. Thus, when $q = 1$, $p = \$399$, and so on. Now try to figure out the value your customer attaches to each of the first two units of your program. Assuming that the customer is rational, the customer will buy a unit if and only if the value of the unit exceeds your price. At a price of $400, demand equals 0, so the first unit cannot be worth $400. At a price of $399, however, demand equals 1 unit. Therefore, the first unit must be worth somewhere between $399 and $400. Similarly, at a price of $399, the customer does not purchase the second unit. At a price of $398, however, the customer is purchasing two units, so the customer does purchase the second unit. Therefore, the customer values the second unit somewhere between $399 and $398. The customer's willingness to pay for a unit of a product is often referred to as the unit's *reservation price*.

It can be shown that a good approximation to the value of the $i^{th}$ unit purchased by the customer is the price that makes demand equal to $i - 0.5$. For example, by setting $q$ equal to 0.5, you can find that the value of the first unit is $400 - 0.5 = \$399.50$. Similarly, by setting $q = 1.5$, you can find that the value of the second unit is $400 - 1.5 = \$398.50$.

Suppose the demand curve can be written as $p = D(q)$. In your example this looks like $D(q) = 400 - q$. The reader who knows integral calculus can exactly determine the value a consumer places on the first $n$ items by computing $\int_0^n D(q)dq$. In your example you would find the value of the first two units to be the following:

$$\int_0^2 (400-q)dq = [400q - .5q^2]_0^2 = 800 - 2 = 798$$

This agrees with your approximate method that yields a value of 399.5 + 398.5 = 798 for the first two units.

# Profit Maximizing with Nonlinear Pricing Strategies

Throughout this chapter assume a power company (Atlantis Power and Light, APL for short) wants to determine how to maximize the profit earned from a customer whose demand in kilowatt hours (kwh) for power is given by $q = 20 - 2p$. It costs $2 to produce a unit of power. Your analysis begins by assuming APL will use linear pricing; that is, charging the same price for each unit sold. You will find that with linear pricing the maximum profit that can be obtained is $32. Then you will find the surprising result that proper use of quantity discounts or a two-part tariff doubles APL's profit! The work for this chapter is in the file Powerblockprice.xls. To determine the profit maximizing linear pricing rule, you simply want to maximize $(20 - 2p) \times (p - 2)$. In the oneprice worksheet from the Powerblockprice.xls file, a price of $6 yields a maximum profit of $32 (see Figure 6-1). The Solver model simply chooses a non-negative price (changing the cell that maximizes profit [Cell D12]). Charging $6 per kwh yields a maximum profit of $32.00.

| | C | D | E |
|---|---|---|---|
| 8 | price | $6.00 | |
| 9 | demand | 8 | |
| 10 | unit cost | $2.00 | |
| 11 | | | |
| 12 | profit | $32.00 | |
| 13 | | | |

**Figure 6-1:** Finding the profit maximizing single price strategy

## Optimizing the Standard Quantity Discount

To determine a profit maximizing pricing strategy that uses the standard quantity discount, assume the quantity discount pricing policy is defined as

follows: All units up to a value *CUT* are sold at a high price (*HP*). Recall that *CUT* is simply the cutoff point at which the per unit price is lowered. All other units sell at a lower price (*LP*). Assuming the customer chooses the number of kwh with the highest non-negative surplus, you can use the Evolutionary Solver to determine profit maximizing values of *CUT*, *HP*, and *LP*. The work for this task is shown in sheet `qd` of file `Powerblockprice.xls`. Also, Figure 6-2 shows that the demand curve may be written as $p = 10 - (q/2)$, so the first unit is valued at $10 - (.5/2) = \$9.75$, the second unit is valued at $10 - (1.5/2) = \$9.25$, and so on.

| | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|
| 1 | | | cutoff | 5 | | Units bought | 16 | |
| 2 | cost | | HP | 14.30422 | | Revenue | $ 96.00 | |
| 3 | 2 | | LP | 2.2253 | | Prod Cost | $ 32.00 | |
| 4 | | | | | Max surplus | 0.000611745 | | |
| 5 | Midpoint | Unit | Value | Cum Valu | Price paid | Surplus | | Profit |
| 6 | 0.5 | 1 | 9.75 | 9.75 | 14.30421665 | -4.554216652 | | $ 64.00 |
| 7 | 1.5 | 2 | 9.25 | 19 | 28.6084333 | -9.608433303 | | |
| 8 | 2.5 | 3 | 8.75 | 27.75 | 42.91264995 | -15.16264995 | | |
| 9 | 3.5 | 4 | 8.25 | 36 | 57.21686661 | -21.21686661 | | |
| 10 | 4.5 | 5 | 7.75 | 43.75 | 71.52108326 | -27.77108326 | | |
| 11 | 5.5 | 6 | 7.25 | 51 | 73.74638371 | -22.74638371 | | |
| 12 | 6.5 | 7 | 6.75 | 57.75 | 75.97168417 | -18.22168417 | | |
| 13 | 7.5 | 8 | 6.25 | 64 | 78.19698462 | -14.19698462 | | |
| 14 | 8.5 | 9 | 5.75 | 69.75 | 80.42228508 | -10.67228508 | | |
| 15 | 9.5 | 10 | 5.25 | 75 | 82.64758553 | -7.647585529 | | |
| 16 | 10.5 | 11 | 4.75 | 79.75 | 84.87288598 | -5.122885984 | | |
| 17 | 11.5 | 12 | 4.25 | 84 | 87.09818644 | -3.098186438 | | |
| 18 | 12.5 | 13 | 3.75 | 87.75 | 89.32348689 | -1.573486892 | | |
| 19 | 13.5 | 14 | 3.25 | 91 | 91.54878735 | -0.548787347 | | |
| 20 | 14.5 | 15 | 2.75 | 93.75 | 93.7740878 | -0.024087801 | | |
| 21 | 15.5 | 16 | 2.25 | 96 | 95.99938826 | 0.000611745 | | |
| 22 | 16.5 | 17 | 1.75 | 97.75 | 98.22468871 | -0.47468871 | | |
| 23 | 17.5 | 18 | 1.25 | 99 | 100.4499892 | -1.449989164 | | |
| 24 | 18.5 | 19 | 0.75 | 99.75 | 102.6752896 | -2.925289618 | | |
| 25 | 19.5 | 20 | 0.25 | 100 | 104.9005901 | -4.900590073 | | |

**Figure 6-2:** Finding the profit maximizing standard quantity discount strategy

To complete the determination of the profit maximizing standard quantity discount strategy, complete the following steps:

1. Copy the formula `=10-0.5*C6` from E6 to E7:E25 to determine the value of each unit.

2. In column F compute the cumulative value associated with buying 1, 2,...19 units. In F6 compute the value of the first unit with formula `=E6`. Copy the formula `=F6+E7` from F7 to F8:F25 to compute the cumulative value of buying 2, 3,...20 units.

3. Copy the formula `=IF(D6<=cutoff,HP*D6,HP*cutoff+LP*(D6-cutoff))` from G6 to G7:G25 to compute the total cost the consumer pays for purchasing each number of units. If you want to analyze different nonlinear pricing strategies, just change this column to a formula that computes the price the customer is charged for each number of units purchased.

4. Copy the formula `=F6-G6` from H6 to H7:H25 to compute the consumer surplus associated with each purchase quantity. In cell H4 compute the maximum surplus with the formula `=IF(MAX(H6:H25)>= 0,MAX(H6:H25),0)`; if it is negative no units will be bought, and in this case you enter a surplus of 0.

5. In cell I1 use the match function to compute the number of units bought with the function `=IF(H4>0,MATCH(H4,H6:H25,0),0)`.

6. In cell I2 compute the sales revenue with the formula `=IF(I1=0,0,VLOOKUP (I1,lookup,4))`. The range D5:G25 is named Lookup.

7. In cell I3 compute production cost with the formula `=I1*C3`.

8. In cell J6 compute profit with the formula `=I2-I3`.

9. Use Solver to find the values of *CUT*, *HP*, and *LP* that maximize profit. The Solver window is shown in Figure 6-3.

You can constrain the value of *CUT* and each price to be at most $20. To do so, charge $14.30 per unit for the first 5 units and $2.22 for remaining units. A total profit of $64 is earned. The quantity discount pricing structure here is an incentive for the customer to purchase 16 units, whereas linear pricing results in the customer purchasing only 6 units. After all, how can the customer stop at 6 units when later units are so inexpensive?

## Optimizing the Nonstandard Quantity Discounts

Now assume that your quantity discount strategy is to charge a high price (*HP*) if at most *CUT* units are purchased, and if more than *CUT* units are purchased, all units are sold at a low price (LP). The only change needed to solve for the profit maximizing strategy of this form (see sheet `qd2`) is to copy the formula `=IF(D6<=cutoff,HP*D6,LP*D6)` from G6 to G7:G25. Then just rerun the Solver window, as shown in Figure 6-3. This pricing strategy, shown in Figure 6-4, maximizes the profit: If the customer buys up to 15 units, he pays $16.79, and if he buys at least 16 units, he pays $6 per unit. Then the customer buys 16 units, and you make $64 in profit.
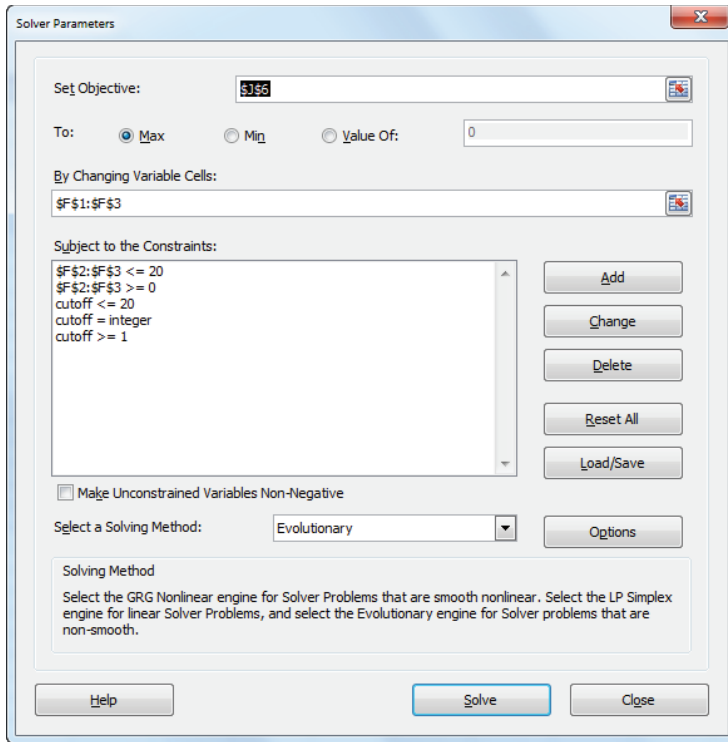
**Figure 6-3:** Determining the profit maximizing standard quantity discount

| | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|
| 1 | | | cutoff | 15 | | Units bought | 16 | |
| 2 | cost | | HP | 16.79609 | | Revenue | $ 96.00 | |
| 3 | | 2 | LP | 5.999873 | | Prod Cost | $ 32.00 | |
| 4 | | | | | Max surplus | 0.002031561 | | |
| 5 | Midpoint | Unit | Value | Cum Valu | Price paid | Surplus | | Profit |
| 6 | 0.5 | 1 | 9.75 | 9.75 | 16.79608788 | -7.046087876 | | $ 64.00 |
| 7 | 1.5 | 2 | 9.25 | 19 | 33.59217575 | -14.59217575 | | |
| 8 | 2.5 | 3 | 8.75 | 27.75 | 50.38826363 | -22.63826363 | | |
| 9 | 3.5 | 4 | 8.25 | 36 | 67.1843515 | -31.1843515 | | |
| 10 | 4.5 | 5 | 7.75 | 43.75 | 83.98043938 | -40.23043938 | | |
| 11 | 5.5 | 6 | 7.25 | 51 | 100.7765273 | -49.77652726 | | |
| 12 | 6.5 | 7 | 6.75 | 57.75 | 117.5726151 | -59.82261513 | | |
| 13 | 7.5 | 8 | 6.25 | 64 | 134.368703 | -70.36870301 | | |
| 14 | 8.5 | 9 | 5.75 | 69.75 | 151.1647909 | -81.41479088 | | |
| 15 | 9.5 | 10 | 5.25 | 75 | 167.9608788 | -92.96087876 | | |
| 16 | 10.5 | 11 | 4.75 | 79.75 | 184.7569666 | -105.0069666 | | |
| 17 | 11.5 | 12 | 4.25 | 84 | 201.5530545 | -117.5530545 | | |
| 18 | 12.5 | 13 | 3.75 | 87.75 | 218.3491424 | -130.5991424 | | |
| 19 | 13.5 | 14 | 3.25 | 91 | 235.1452303 | -144.1452303 | | |
| 20 | 14.5 | 15 | 2.75 | 93.75 | 251.9413181 | -158.1913181 | | |
| 21 | 15.5 | 16 | 2.25 | 96 | 95.99796844 | 0.002031561 | | |
| 22 | 16.5 | 17 | 1.75 | 97.75 | 101.9978415 | -4.247841466 | | |
| 23 | 17.5 | 18 | 1.25 | 99 | 107.9977145 | -8.997714493 | | |
| 24 | 18.5 | 19 | 0.75 | 99.75 | 113.9975875 | -14.24758752 | | |
| 25 | 19.5 | 20 | 0.25 | 100 | 119.9974605 | -19.99746055 | | |

**Figure 6-4:** Profit maximizing nonstandard quantity discount

## Optimizing a Two-Part Tariff

Organizations such as a golf country club or Sam's Club find it convenient to give the consumer an incentive to purchase many units of a product by charging a fixed fee to purchase any number of units combined with a low constant per unit cost. This method of nonlinear pricing is called a *two-part tariff.* To justify paying the fixed fee the customer needs to buy many units. Of course, the customer will not buy any units whose unit cost exceeds the consumer's reservation price, so the organization must be careful to not charge a unit price that is too high.

For example, when you join a country club, you often pay a fixed fee for joining and then a constant fee per round of golf. You might pay a $500 per year fixed fee to be a club member and also pay $30 per round of golf. Therefore one round of golf would cost $530 and two rounds of golf would cost $560. If a two-part tariff were linear then for any number of purchased units, the cost per unit would be the same, but because one round cost $530 per round and two rounds costs $530/2 = $265, the two-part tariff is not consistent with linear pricing.

To dig a little deeper into the concept of a two-part tariff, look at the worksheet tpt. You will see how a power company can optimize a two-part tariff by performing the following steps:

1. To begin, make a copy of your QD or QD2 worksheet.
2. In F2 enter a trial value for the fixed fee, and in F3 enter a trial value for the price per unit. Name cell F2 Fixed and cell F3 var. The only formulas that you need to change are the formulas in the Price Paid formula in Column G (see Figure 6-5).
3. Copy the formula =Fixed+D6*Var from G6 to G7:G25, to compute the cost of buying each number of units.
4. Adjust the Solver window, as shown in Figure 6-6.

You need to constrain the fixed and variable costs between $0 and $100. The profit is maximized by charging a fixed fee of $60.27 and a cost of $2.21 per unit. Again a maximum profit of $64 is made, and the customer purchases 16 units.

| ▲ | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|
| 1 | | | | 0 | | Units bought | 16 | |
| 2 | cost | | Fixed | 60.57433887 | | Revenue | $ 96.00 | |
| 3 | 2 | | var | 2.214067061 | | Prod Cost | $ 32.00 | |
| 4 | | | | | Max surplus | 0.000588158 | | |
| 5 | Midpoint | Unit | Value | Cum Value | Price paid | Surplus | | Profit |
| 6 | 0.5 | 1 | 9.75 | 9.75 | 62.78840593 | -53.03840593 | | $ 64.00 |
| 7 | 1.5 | 2 | 9.25 | 19 | 65.00247299 | -46.00247299 | | |
| 8 | 2.5 | 3 | 8.75 | 27.75 | 67.21654005 | -39.46654005 | | |
| 9 | 3.5 | 4 | 8.25 | 36 | 69.43060711 | -33.43060711 | | |
| 10 | 4.5 | 5 | 7.75 | 43.75 | 71.64467417 | -27.89467417 | | |
| 11 | 5.5 | 6 | 7.25 | 51 | 73.85874123 | -22.85874123 | | |
| 12 | 6.5 | 7 | 6.75 | 57.75 | 76.07280829 | -18.32280829 | | |
| 13 | 7.5 | 8 | 6.25 | 64 | 78.28687535 | -14.28687535 | | |
| 14 | 8.5 | 9 | 5.75 | 69.75 | 80.50094241 | -10.75094241 | | |
| 15 | 9.5 | 10 | 5.25 | 75 | 82.71500948 | -7.715009476 | | |
| 16 | 10.5 | 11 | 4.75 | 79.75 | 84.92907654 | -5.179076537 | | |
| 17 | 11.5 | 12 | 4.25 | 84 | 87.1431436 | -3.143143598 | | |
| 18 | 12.5 | 13 | 3.75 | 87.75 | 89.35721066 | -1.607210659 | | |
| 19 | 13.5 | 14 | 3.25 | 91 | 91.57127772 | -0.57127772 | | |
| 20 | 14.5 | 15 | 2.75 | 93.75 | 93.78534478 | -0.035344781 | | |
| 21 | 15.5 | 16 | 2.25 | 96 | 95.99941184 | 0.000588158 | | |
| 22 | 16.5 | 17 | 1.75 | 97.75 | 98.2134789 | -0.463478903 | | |
| 23 | 17.5 | 18 | 1.25 | 99 | 100.427546 | -1.427545964 | | |
| 24 | 18.5 | 19 | 0.75 | 99.75 | 102.641613 | -2.891613025 | | |
| 25 | 19.5 | 20 | 0.25 | 100 | 104.8556801 | -4.855680086 | | |

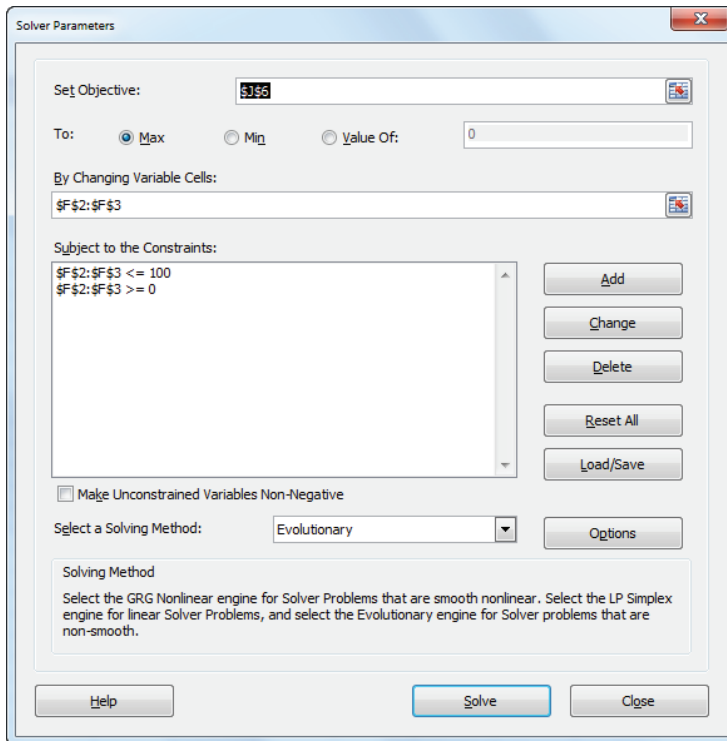**Figure 6-5:** Optimizing a two-part tariff



**Figure 6-6:** Solver window for two-part tariff

All three of the fairly simple nonlinear pricing strategies result in doubling the linear price profit to $64. There is no nonlinear pricing plan that can make more than $64 off the customer. You can easily show the exact value of the sixteenth unit is $2.25, and the exact value of the seventeenth unit is $1.75, which is less than the unit cost of $2.00. Because the customer values unit 16 more than the cost, you should be able to get her to buy 16 units. Because the area under the demand curve (see Exercise 5) from 0 to 16 is 96, the most you can make the consumer pay for 16 units is $96. Therefore your maximum profit is $96 − ($2)(16) = $64, which is just what you obtained using all three nonlinear pricing strategies. The problem with the linear pricing strategy is that it does not extract the surplus value in excess of the $2 cost that the customer places on earlier units. The nonlinear pricing strategies manage to extract all available consumer surpluses.

For a monopolist (like a power company) it may be realistic to assume that an organization can extract all consumer surplus. In the presence of competition, it may be unrealistic to assume that an organization can extract all consumer surplus. To illustrate how the models of this chapter can be modified in the presence of competition, suppose you are planning to enter a business where there is already a single company selling a product. Assume that this company is currently extracting 80 percent of the consumer surplus and leaving 20 percent of the consumer surplus with the consumer. By adding a constraint to your models that leaves, say 25 percent of the consumer surplus in the hands of the customer, you can derive a reasonable nonlinear pricing strategy that undercuts the competition. For an illustration of this idea see Exercise 8.

## Summary

In this chapter you learned the following:

- Nonlinear pricing strategies involve not charging the same price for each unit sold.
- If you assume customers purchase the available option with the largest (if non-negative) surplus, then you can use the Evolutionary Solver to find profit maximizing nonlinear pricing strategies such as quantity discounts and two-part tariffs.
- For a linear demand curve, nonlinear pricing earns twice as much profit as linear pricing.

## Exercises

1. You own a small country club. You have three types of customers who value each round of golf they play during a month as shown in the following table:

| Round | Segment 1 | Segment 2 | Segment 3 |
|---|---|---|---|
| 1 | $60 | $60 | $40 |
| 2 | $50 | $45 | $30 |
| 3 | $40 | $30 | $20 |
| 4 | $30 | $15 | $10 |
| 5 | $20 | $0 | $0 |
| 6 | $10 | $0 | $0 |

It costs you $5 in variable costs to provide a customer with a round of golf. Find a profit maximizing a two-part tariff. Each market segment is of equal size.

2. The demand curve for a product is $q = 4000 - 40p$. It costs $5 to produce each unit.

   a. If you charge a single price for each unit, how can you maximize profit?
   b. If you use a two-part tariff, how can you maximize the profit?

3. Using the data in Exercise 1, determine the profit maximizing the quantity discount scheme in which all units up to a point sell for a high price, and all remaining units sell for a low price.

4. The file Finalmusicdata.xls (available on the companion website) contains data showing the most 1,000 people were willing to pay for 10 songs. For example, as shown in Figure 6-7, Person 1 is willing to pay up to 38 cents for Song 1.

| | C | D | E | F | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | | Song 1 | Song 2 | Song 3 | Song 4 | Song 5 | Song 6 | Song 7 | Song 8 | Song 9 | Song 10 |
| 4 | 1 | $0.38 | $0.94 | $1.32 | $1.85 | $0.57 | $1.16 | $1.39 | $2.23 | $1.95 | $1.88 |
| 5 | 2 | $1.02 | $1.83 | $2.23 | $0.56 | $0.30 | $2.12 | $2.09 | $1.63 | $1.65 | $1.50 |
| 6 | 3 | $0.60 | $1.44 | $1.91 | $1.67 | $0.82 | $2.41 | $0.64 | $1.00 | $1.80 | $1.46 |
| 7 | 4 | $0.31 | $2.10 | $1.12 | $0.84 | $1.30 | $0.61 | $0.84 | $1.95 | $0.88 | $0.51 |
| 8 | 5 | $1.71 | $1.22 | $1.76 | $2.35 | $0.99 | $1.96 | $2.04 | $0.65 | $2.33 | $1.56 |
| 9 | 6 | $1.44 | $2.35 | $2.27 | $0.94 | $1.20 | $0.43 | $0.51 | $1.97 | $1.85 | $1.51 |
| 10 | 7 | $0.22 | $1.04 | $1.77 | $1.51 | $0.91 | $2.11 | $1.33 | $2.10 | $2.14 | $0.61 |
| 11 | 8 | $1.27 | $0.93 | $2.40 | $1.57 | $1.91 | $0.35 | $1.14 | $1.24 | $1.70 | $0.63 |
| 12 | 9 | $1.85 | $0.36 | $0.28 | $0.61 | $1.75 | $1.59 | $1.55 | $1.00 | $0.43 | $1.37 |
| 13 | 10 | $0.85 | $2.12 | $1.29 | $2.41 | $1.02 | $0.42 | $1.18 | $2.32 | $1.27 | $0.58 |
| 14 | 11 | $2.07 | $2.24 | $1.26 | $1.49 | $0.25 | $0.42 | $0.75 | $2.09 | $1.65 | $1.14 |

**Figure 6-7:** Data for Problem 4

a. If you charge a single price for each song, what price can maximize your revenue?

b. If you use a two-part tariff pricing scheme (for example $3 to download any songs + 40 cents per song), how can you maximize revenue?

c. How much more money do you make with the two-part tariff?

**NOTE**  For the second part, you must realize that a person will buy either nothing, their top-valued song, their top 2 songs, their top 3 songs,…their top 9 songs or top 10 songs.

5. Show that for the demand curve $q = 40 - 2p$ the consumer places a value of exactly $256 on the first 16 units.

6. (Requires calculus) Suppose the demand for a computer by a leading corporate client (for $p \geq 100$) is given by $200000 / p^2$. Assume the cost of producing a computer is $100.

a. If the same price is charged for each computer, what price maximizes profit?

b. If you charge HP for the first CUT computers and LP for remaining computers purchased, how can you maximize profit?

7. Suppose Verizon has only three cell phone customers. The demand curve for each customer's monthly demand (in hours) is shown here:

| Customer | Demand curve |
|---|---|
| 1 | $Q = 60 - 20p$ |
| 2 | $Q = 70 - 30p$ |
| 3 | $Q = 50 - 8p$ |

Here $p$ = price in dollars charged for each hour of cell phone usage. It costs Verizon 25 cents to provide each hour of cell phone usage.

a. If Verizon charges the same price for each hour of cell phone usage, what price should they charge?

b. Find the profit maximizing a two-part tariff for Verizon. How much does the best two-part tariff increase the profit over the profit maximizing the single price?

8. For the power company example construct a standard quantity discount pricing strategy that leaves the customer with 25 percent of the consumer surplus.