# Table of Contents

Dedications ................................................................. viii

About the Authors ....................................................... viii
  Oz du Soleil ..................................................................... viii
  Bill Jelen ........................................................................ x

Acknowledgements ....................................................... x

Introduction: Welcome to the World of Guerrilla Data Analysis! ........................................... xii
  In The Heat of Conflict .................................................. xii
  Small, Stupid Stuff ....................................................... xii
  Big, Complicated Stuff .................................................. xiii

How to Use This Book .................................................... xiv

Download Files .............................................................. xiv

Reviewing the Basics ..................................................... 1
  Changing Formulas to Values ....................................... 1
  Using Paste Special in Other Ways ................................ 3
    Transposing Columns and Rows ................................ 3
    Performing a Calculation on Every Cell in a Range .... 4
  Using Helper Columns ................................................ 6
  Using Relative, Absolute, and Mixed References .......... 8

Developing Dynamic Spreadsheets ................................. 10

Conditional Formatting ................................................ 12
  Using Conditional Formatting to Make Deadline Alerts .... 13
  Using Conditional Formatting to Find Duplicates ............ 16
  Using Icons with Conditional Formatting ..................... 17

Using IF Statements ...................................................... 20
  Using an IF Statement with COUNTIF ........................... 21

Sorting ........................................................................ 22
  Rules for Sorting ........................................................ 22
    The Data Range Must Be Contiguous (No Completely Blank Rows or Columns) .... 22
    Headings Must Be Only One Cell Tall ............................ 24
    Merged Cells Do Not Sort ........................................... 25
  Using Sorting ............................................................ 26
    Using the Quick-Sort Icons ........................................ 26
    Understanding Ascending and Descending Sorts .......... 27
    Let’s Get Some Data Sorted Out ................................. 28
    Sorting with Two Criteria .......................................... 29
    Sorting with the Help of a Helper Column .................... 30

Filtering ................................................................... 32
  What You Need to Know About Filtering Before You Do It ... 32
  Getting Down to Filtering Business ............................... 33
    What’s In the Data Set? .............................................. 34
  Finding Records Quickly with AutoFilter ....................... 36
  Filtering to Find the Top (or Bottom) Five Transactions .... 38
Using Pivot Tables
.............................................................................................. 76
What Is a Pivot Table? ........................................................................ 76
  Example 1: Summing Values with a Pivot Table................................. 77
  Example 2: Counting Values with a Pivot Table ............................... 77
  Example 3: Filtering with a Pivot Table ........................................... 78
  Example 4: Using a Pivot Table to Find a Sum and an Average at the Same Time .......................................................... 79
Creating a Pivot Table......................................................................... 79
  Summing Values with the Pivot Table ............................................. 81
  Counting Values with the Pivot Table ............................................. 83
  Filtering with the Pivot Table ....................................................... 84
  Grouping Dates in the Pivot Table ................................................ 85
  Using the Pivot Table to Get a Sum & Average at the Same Time ....... 87
  Using the Pivot Table to Get the Percentage of the Total ................. 87
  Using the Pivot Table to Filter for the Top Five .............................. 89
  Using the Pivot Table to Drill Down for Isolated Details .................. 90
Deleting a Pivot Table ........................................................................ 91
Overriding the Default Row Sequence in a Pivot Table ....................... 91
Rearranging Pivot Table Headers ...................................................... 92
# TABLE OF CONTENTS

Using Calculated Items & Calculated Fields ............................................................... 93
  Working with Calculated Items .............................................................................. 93
  Working with Calculated Fields .......................................................................... 96
  Final Notes on Calculated Items and Calculated Fields ........................................ 97
Pivot Table Q&A ........................................................................................................ 98
Pivot Table Conclusions .......................................................................................... 100

**Using Array Formulas** .......................................................................................... 101
  Basic Array Formula ............................................................................................... 101
  You Cannot Change Part of an Array .................................................................... 102
  Copying an Array Formula ...................................................................................... 103
  Modifying an Array Formula ................................................................................. 103
  Using FREQUENCY to Create a Histogram ............................................................ 103
  Going One Step Further: An Array Inside an Array Formula ................................. 108
  Array Formulas and System Memory ...................................................................... 109

**Stepping Up to Excel Tables** ............................................................................... 109
  Converting a Data Range to a Table ....................................................................... 109
  Adding New Data to a Table ................................................................................... 112
  Adding a New Column ............................................................................................ 113
  Using SUMIFS with Table References .................................................................. 113
  Using the Table Design Tab ..................................................................................... 114
  Other Cool Table-Related Stuff .............................................................................. 115
  Some Warnings About Working With Tables ........................................................ 118
    Mixing Formulas in a Column .............................................................................. 118
    Adding New Data to a Table ................................................................................ 118
    Sheet Protection: Tables Must Be Completely Protect or Completely Unprotected ...... 118
  Excel Tables Conclusion ........................................................................................ 119

**Using the INDIRECT and OFFSET Functions** ....................................................... 119
  Using INDIRECT ....................................................................................................... 119
  Using INDIRECT with VLOOKUP ........................................................................... 121
  Using INDIRECT in an Array Formula ..................................................................... 121
  Using OFFSET ......................................................................................................... 122

**Controlling Data Inputs and Maintaining Data Integrity** .................................. 124
  Data Validation Overview ...................................................................................... 124
  Implementing Dropdown Lists ............................................................................... 125
  Controlling Dates ..................................................................................................... 126
  Ensuring Reasonable Numbers .............................................................................. 127
  Preventing Start and End Times from Being Reversed .......................................... 128
  Data Validation Conclusions .................................................................................. 129

**Implementing Error-Handling and Formula Triggers** .......................................... 130
  Error-Handling Example ....................................................................................... 130
  Formula Trigger Example ...................................................................................... 130
  Error-Handling Functions: IFNA vs. IFERROR ....................................................... 131

**Using Pivot Charts** ............................................................................................. 132
  Playing Around with a Pivot Chart ....................................................................... 135
  Adding New Data to a Pivot Chart ....................................................................... 136
  Changing the Chart Type ....................................................................................... 137
Using Slicers with Pivot Charts ................................................................. 138
Using Slicers ......................................................................................... 139
Excel 2013: Guerrilla Data Analysis Gets Real ....................................... 141
   Using Slicers with Tables in Excel 2013 .......................................... 141
   Understanding Data Models and Relationships ............................... 143
Graphing and Charting ........................................................................ 147
Using Excel on the Internet ................................................................. 155
   Embedding Excel in a Blog Post ..................................................... 155
   Differences between Excel Web App and Desktop Excel ............... 158
Down and Dirty Tips and Insights ......................................................... 159
   Overview of Excel Formulas and Functions .................................... 159
   Forcing a Report to Fit on One Page ............................................. 161
   Boolean Logic ............................................................................... 161
   Using & to Mix Text and Numbers in the Same Cell ..................... 162
   Using Text-Manipulation Functions ............................................. 163
   Handling Dates ............................................................................ 163
   Handling Time ............................................................................ 167
   Converting All Results to Minutes .............................................. 170
Useful Excel Functions ......................................................................... 171
   Using PMT to Predict a Loan Payment .......................................... 171
   Using FORECAST ......................................................................... 171
   Using RANK .............................................................................. 172
   Using RANK, RANK.EQ, and RANK.AVG ................................. 172
   Breaking Ties Based on Position in a List .................................... 173
   Using CEILING and FLOOR ....................................................... 175
   Using MAX, MIN, LARGE, and SMALL ..................................... 175
   Using COUNTA ........................................................................... 176
   Using CONVERT ......................................................................... 176
   Using ABS to Compare Errors in Absolute Terms ....................... 176
   Using RAND and RANDBETWEEN .......................................... 177
   Using RANDBETWEEN and INDEX to Make Random Assignments 178
   Using CHOOSE ............................................................................ 179
   Putting CHOOSE to Work ............................................................ 180
   Using SUMPRODUCT ..................................................................... 181
   Using EOMONTH ........................................................................ 181
Troubleshooting Excel .......................................................................... 182
   Quickly Checking Sums and Averages ........................................ 182
   Volatile, Slow, and Peculiar Functions and Features ...................... 184
   Using SUM vs. Adding Individual Cells ...................................... 185
   Troubleshooting by Using CTRL+` (also known as Ctrl+~) ............ 186
      Crossfooting ............................................................................ 187
   Using the Formula Evaluator ....................................................... 187
   Troubleshooting by Checking Highlighted Ranges in a Formula ...... 189
   Unhiding Column A .................................................................... 190
   Getting Rid of Gridlines .............................................................. 190
Linking ................................................................................................. 190
   Linking Between Worksheets ...................................................... 190
TABLE OF CONTENTS

Linking One Cell to Another on a Different Worksheet ..................................................... 192
Using a VLOOKUP with References to Another Worksheet ............................................... 193
Inserting Table References Between Worksheets .............................................................. 193
Linking Workbooks ............................................................................................................ 194

Spreadsheet Layout and Development .............................................................................. 195
Digging into the Details of the Layout ................................................................................ 196
Taking Advantage of Good Spreadsheet Layout ................................................................. 199
Adding Some Data and Making a Pivot Table ..................................................................... 200

Using Keyboard Shortcuts ................................................................................................. 202
Quickly Navigating Using the Ctrl or End Key ................................................................. 202
Navigating Between Worksheets ....................................................................................... 203
Formatting Shortcuts ......................................................................................................... 203
Clipboard Shortcuts ........................................................................................................... 203
Calculation Shortcuts ......................................................................................................... 204
Editing Shortcuts ................................................................................................................ 204
Excel Commands................................................................................................................ 204
F4 Repeats Last Command ................................................................................................. 205
The Amazing F4 Key (While Editing a Formula) ................................................................. 205
Using Arrow Keys to Enter a Formula ................................................................................ 205

Index .................................................................................................................................. 207
Dedications

To my mother, Maere Floyd, who seemed like the meanest mother in the world. I’d ask her a question like, “Who was the president when I was born?” Her response would be, “Did you try to look it up?” Eventually, I learned that the only right response was “Yes.” I had to have at least tried.

We had a World Book Encyclopedia set in the living room, and those books were meant for use, not decoration.

All these many years later, one lesson has stuck with me and has served me as an analyst, consultant, and Excel power user: If there’s a fact in the world, make an effort to look for it. Don’t make stuff up. Dig, probe, and pick until you get your answer.

As a high school teacher’s aide preparing for retirement, my mother has remained committed to young folks and how they learn. Thanks, Mom!

—Oz du Soleil

To Dan Bricklin and Bob Frankston.

—Bill Jelen

About the Authors

Oz du Soleil

Biographies tend to be written in the third-person, but let me speak directly to you.

I grew up in the city of North Chicago, Illinois. I joined the US Navy in 1985 and spent six years there. In that time, I served on the fast attack submarine USS Sturgeon (SSN-637) and the fast frigate USS Joseph Hewes (FF-1078).

After the Navy, I graduated from University of Illinois at Chicago (UIC) with a BA in philosophy and a minor in economics and two years of Russian language study. My coursework in philosophy involved logic and decision theory, which offers some insight into my passion for Excel and data—lots of strategy, symbols, variables, and parentheses.

After UIC, I spent two years at Harvard’s Kennedy School of Government. I didn’t finish the master’s in public administration. Public policy just wasn’t my thing, although while I was there, I did help with legal research for the New York Civil Liberties Union in a successful case against Time-Warner Cable.
Then I was back to Chicago—and Long Beach, California; and Maplewood, New Jersey; and New York City; and Chicago again where I eventually ended up in a job where I used Excel to investigate complaints. Duplicate orders were shipping out, test results weren’t showing up, accounts were assigned to the wrong sales reps, the wrong people were getting termination warning letters. In all those cases, the culprit turned out to be bad data. Some of those problems weren’t mine to solve, but management allowed me to dig into the data, get an understanding of the processes, work with the stakeholders, clean the data, and revise the processes. I loved putting out those fires!

For seven years I put out such fires: uncovering flaws in reports, exposing how calculations in reports didn’t match business rules, modifying processes, cleaning data, and generally reducing the unnecessary misery caused by crap data. And then, at a mandatory meeting in January 2008, it was announced that the company was moving to LaCrosse, Wisconsin, and all 300 of us would be laid off over the next 12 months. I was laid off in July 2008 and decided to freelance my Excel skills.

While I took on Excel clients and taught workshops, I also founded a nonprofit and was its unpaid executive director for three years. And I developed a ghost pepper peanut butter cookie that sold quite successfully.

Today I run a blog, DataScopic.net, that focuses on data and data management, with Excel at the root of it all. After seeing so much data in my own jobs, in working with clients and students, and in discussions with other data folks, my perspective is that Excel is just a tool. We need data-savvy people and solid processes—in addition to the right tools—in order to keep our data clean...and keep termination warning letters from going to the wrong people.

I love freelancing and taking projects, but I love teaching and speaking even more. That led me to take improv and stand-up comedy courses at Second City. It’s been great! They’re wonderful for anyone who regularly gets up in front of people.

Here are a few more things I’d like to say about myself:

- I always carry Sriracha with me. (Gotta be prepared if friends want to go eat Italian or some other cuisine void of capsaicin.)
- I wrestled for 13 years—from junior high school, high school, college and a few Navy tournaments.
- I prefer bow ties and custom-made hats.
- I’ve visited Gdansk, Poland, and Budapest, Hungary, and I got tattoos in both places.
- I am in a film called *Sriracha: A Documentary Film by Griffin Hammond*.
- The writing of this book was briefly interrupted by my move from Chicago, IL to Portland, OR.

Thanks for buying this book! I welcome your feedback, at Oz@DataScopic.net.
Bill Jelen

Bill Jelen wrote the first edition of this book in 2002. Since then, he has written 42 other books about Microsoft Excel and does his Power Excel seminar anywhere he can find a room full of accountants. He has run the MrExcel.com website since 1998.

Acknowledgements

Thanks to Bill Jelen, Mr. Excel, for the opportunity to write this book, and his leadership in the Excel community.

Thanks to Troy Berry for seeing that I was an analyst before I thought of myself as an analyst. Through his prodding and advice, I began to see this Excel thing as a powerful tool—far more than a fun toy for solving problems.

Thanks to Helena Bouchez for her real-deal coaching that kept me on track and got this book done on time!

Thanks to Keidra Chaney for many many conversations about data and the needs that non-data people have as they are thrust into data-driven roles. Thanks for your shared commitment in empowering those people in understanding the strengths and limitations of data.

Thanks to Monica Johnson, Lupe Miranda, Drew Alexander, and Nancy Migalla. These former co-workers were incredibly supportive not only as co-workers but as teachers who showed me the world of Pivot Tables, the importance of data quality, and that being an analyst goes beyond the data. It includes managing people who don’t like the results and owning up to it when someone finds mistakes in the analysis.

Thanks to Mike “ExcellIsFun” Girvin, Chandoo, Jocelyn and Rob Collie, Szilvia Juhasz, Krisztina Szabó, Maxime Emanuel, Kevin Lehrbass, Craig Hatmaker, Petros Chatzipantazis, Rahim Zulfiqar Ali, Mynda Treacy, Ann Emery, Hiran De Silva, and countless others who help make the Excel community supportive, vibrant, and international—and who remind us why Excel is the number-one business intelligence tool in the world.

Thanks to my fellow co-hosts of Excel TV, Rick Grantham and Jordan Goldmeier, and our wonderful Excel TV guests: Zach Barresse, Keidra Chaney, Bill Jelen, Szilvia Juhasz, Kari Finn, Chandoo, Jon Peltier, Mynda Treacy, Johann Odou, John Persico, Mike “ExcellIsFun” Girvin. We’re bringing knowledge, a wide range of insights, fire, and wild adventure!
Thanks to supporters Patrick Richards, Jeff Bradford, Tshombe Brown, Charlie Vlahogiannis, Kate Christenson, the EBassist community, Heather Dart, Dianna Smith, Andreas Pavlatos and too many others to name. I appreciate your passion and interest.

Thanks to Jeff Ashton, the first person who ever told me, “No, there isn’t an Excel feature for that, but here’s a way we can trick Excel into doing it anyway.” To me, that’s when you’re really using Excel—when you’re putting the right formulas and features together to get your work done.

Thanks to Ranjit Souri, my stand-up comedy instructor at Second City, who warned, “You will bomb! You must bomb!” Those eight weeks in Ranjit’s class were a solid lesson that failure is part of any learning process and that you can fail in front of an audience—as long as you’re learning and putting in the work to get better. Analysts do fail in front of others. Good analysts own their mistakes. Bad analysts end up fired or jailed.

Thanks to Andy Crestodina, Amanda Gant, and the Orbiteers Wine & Web events that encourage good, sincere web content. Those events helped give me the opportunity to write this book. In the early days of my blog, when there were no visitors, there was Andy’s quote: “Crickets never hurt anybody.”

Thanks to my clients, who’ve provided the opportunity to help with their data and challenged me with unique needs—stretching my skill and testing Excel’s limits.

Thanks to my workshop students—both public and private—for asking tough questions, being committed to quality data, and wanting to get more out of Excel. And thanks for the feedback that helps me to improve as an instructor.

—Oz

Thanks to Oz for rescuing this second edition of Guerilla Data Analysis from the long list of “I will have to do that someday” and making it a reality. When I started reading Oz’s blog and seeing him on Excel TV, it struck me that he was the right guy to carry on the Guerilla Data Analysis franchise.

Thanks to everyone in “row 2” of my seminars.

Thanks to David Gainer, Sam Radakovitz, Ben Rampson, Dan Battigan, Joe Chirilov, Joe Camp, and Kari Finn at Microsoft.

Thanks to Mary Ellen Jelen.

—Bill Jelen
Introduction: Welcome to the World of Guerrilla Data Analysis!

Over the years that I’ve been consulting, teaching workshops, and writing a blog, it’s gotten clear to me that there are a lot of people who are in data-driven roles but don’t have a data background. They aren’t sure what Excel can really do, but spreadsheets keep showing up in their inboxes. One of my students complained that she got a promotion, more money, and the title Social Media Strategist. However, instead of getting more social media activity, she got a mountain of data and was directed to “find something interesting in this.” She had become an unwitting data analyst who didn’t know where to start.

Other students and clients have told stories about taking a week to manually compare lists that were thousands of rows long; retyping data that came to them in ALL CAPS; spending days creating summaries while not knowing that Pivot Tables are designed to make those summaries in seconds. This is the world of guerrilla data analysis, where you’re in the heat of data conflict, without formal training, and you need to make something happen.

If you’re reading this, you’re probably a guerrilla analyst, and hopefully you’ll get useful tips and insights from this book, as well as solutions that end unnecessary misery. The examples here are practical and cover a wide variety of areas, including nonprofits, web analytics, accounting, cooking, and retail. The goal is to get your ideas churning by exposing you to the variety of ways to use Excel.

In The Heat of Conflict

Data is coming from everywhere, about all kinds of things. All the cool kids are talking about big data, data science, and predictive analysis. But data also presents everyday problems that aren’t as glamorous as millions of rows of data and Nate Silver coming to tell us the future. Data analysis involves both the small, stupid stuff and big, complicated stuff.

Small, Stupid Stuff

One afternoon in 2005, I needed to print certificates on expensive paper. It was late in the afternoon, and the certificates absolutely had to ship that day. I did the Excel–Word mail merge, and the certificates were coming out of the printer with weirdo numbers instead of dates: 38491, 38464, and 38478 instead of 5/19/2005, 4/22/2005, and 5/6/2005. C’mon! Really?! Now?!
It took me two hours to learn that the Short Date formatting in Excel had been changed to General, and I only had to change it back and redo the merge. But the glitch had already cost me at least 50 sheets of fancy paper, the afternoon was gone, and I’d done a lot of worrying that I’d have to tell a lot of people that the certificates weren’t going to ship on time.

That afternoon I didn’t really have an analysis problem, but my story is a good example of an analyst going into panic mode after all the true hard-core analysis has been done. It’s an example of how small, stupid details can turn a process sideways. Another reason for telling this story is to let you know that you’re not the only one who’s been stopped by small things.

Students ask about these types of disruptions and start their questions with “This may seem like a small thing, but...” It’s not small when you’re under pressure and a whole process has stopped. Guerrilla conflict is guerrilla conflict.

Big, Complicated Stuff

Data analysis involves all kinds of tasks—from investigating known or perceived problems to forecasting to developing dashboards to cleaning data to doing financial modeling—while also satisfying the end users.

One of my first, and favorite, freelancing jobs involved supporting a project manager. She was leading the migration of 400 phone systems (with 200 columns of data for each of the 400 rows) and needed a summary page that provided a simple overview. We spent hours together, going over the complex calculation of contract expiration dates, prorated charges, disconnect charges, disconnect dates, connection dates, additional fees associated with carrier X but not carriers Y and Z, and fees associated with certain systems. This all had to be right so we could show each of the 400 managers what they would save by migrating early rather than migrating at the deadline, 15 months away.

With extremely complicated formulas, having a strategy is key to getting the calculations right. You can do complex calculations in one big formula, but you end up creating a parentheses torture chamber. In devising a calculation strategy, it’s important for an analyst to answer two questions:

- If I have to come back to this in three months, how easily can I figure out what I did?
- If someone else has to figure it out, how easily can that person trace my thinking?

Breaking complicated calculations into pieces, using helper columns, and creating lookup grids will help build those complex calculations and make it easier (for the original analyst or someone else) to troubleshoot later.
How to Use This Book

Excel has many features, and every analyst (and situation) has different needs. So, rather than providing an Excel A-to-Z, this book moves in a rapid-fire way, attempting to plug common holes and give quick exposure to the best of Excel. This book provides real-life examples from a wide variety of areas and presents a few simple examples to isolate certain critical concepts. Most sections are short, but the sections on Pivot Tables and filtering are longer because those topics are so vast.

The book assumes that you have some level of familiarity with Excel. It also assumes that you work with data and need Excel to cooperate. You can read straight through the book, and you can also use it as a reference, when you either have an immediate challenge or simply wonder, “Is there a better way?”

Download Files

If you would like to work along with any example in this book, you can download the data from http://www.datascopic.net/gdav2-downloads
In cell B1 you input the item you’re looking for. In B2 you enter the month that starts a period you’re interested in. In B3 you enter the number of months you want to sum.

You can use the following helper formulas to build the OFFSET function in smaller pieces:

- In C1, =MATCH(B1,F4:F7,FALSE) finds the row number of the item.
- In C2, =MATCH(B2,G3:R3,FALSE) finds the column number of the month.

What if you want to know how many blue shirts were sold in the three months May, June, and July? The formula in cell B6 will give you the answer:

\[
=\text{SUM} (\text{OFFSET} (F3, C1, C2, 1, B4))
\]

Here is a deconstruction of the formula and what it’s doing:

<table>
<thead>
<tr>
<th>Formula Components</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>=SUM(</td>
<td>Opens the SUM function</td>
<td></td>
</tr>
<tr>
<td>OFFSET(</td>
<td>Opens the OFFSET function</td>
<td></td>
</tr>
<tr>
<td>F3</td>
<td>Starts at cell F3</td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>Number of rows</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>Number of columns</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Height of the final range</td>
<td></td>
</tr>
<tr>
<td>B4</td>
<td>Width of the final range</td>
<td></td>
</tr>
<tr>
<td>)</td>
<td>Closes the OFFSET function</td>
<td></td>
</tr>
<tr>
<td>)</td>
<td>Closes the SUM function</td>
<td></td>
</tr>
</tbody>
</table>

The formula tells you that 489 blue shirts sold in May through July.
Controlling Data Inputs and Maintaining Data Integrity

When no effort is made to control user inputs, the results can range from embarrassment to catastrophe. So controlling and maintaining data is an essential topic in guerilla data analysis. Far too often, when the stakes are high, this aspect of spreadsheet development is overlooked. In some cases, too, forensic accountants must try to figure out whether poor spreadsheet integrity was a blunder or a deliberate decision to hide illegal activity in a bunch of confusion.

Controlling and maintaining data is especially a big deal since so many people are working with data and Excel in inventive ways. The upshot is that sloppy spreadsheet use can impact the global economy or merit badges in a Girl Scout troop.

Data Validation Overview

In Excel, data validation is different from validating whether a formula is correct. In this section we’ll look at several examples to see what Excel calls data validation. You can use the data in the figure below to do a lot of things. For one thing, you could calculate the totals charged by each inspector and tally their commissions. But you’ve got several problems:

- In A4, Andrew Kim is called Andy Kim, and Andrew Kim won’t get credit for Andy Kim’s work.
- In A5, Kate Ambrose is called Kate-Marie Ambrose.
- In A7, David has no last name specified.
- In B11, this should probably be a 2015 date, not a 2014 date.

Common input errors such as these prevent you from generating trustworthy results. But you can implement validation to stop some of this nonsense.
Implementing Dropdown Lists

The solution to inconsistent entry of names is to put the choices into dropdown lists. This way, there's only one possible entry for each inspector. Here's what you do:

1. Place your cursor in A2 so you can put the first dropdown list there.
2. Select Data | Data Tools | Data Validation. The Data Validation dialog appears.
3. In the Data Validation dialog, under Allow, select List.
4. Ensure that Ignore Blank is not selected.
5. Ensure that the In-cell Dropdown check box is selected.
6. Click the Edit Reference icon at the far right of the Source field. As shown in the next figure, the Data Validation window minimizes and allows you to select the range where you've stored the dropdown entries.

7. Select the range and click the far-right icon again to bring back the full Data Validation dialog.
8. Click OK.

As shown here, the dropdown list is now in place, and you can copy it down the whole column:
Controlling Dates

If you want to restrict dates to 2015, here’s how you fill out the Data Validation dialog to set up that data validation:

Dates before or after 2015 will result in an error message.
Ensuring Reasonable Numbers

Say that you want to allow a grade to be any positive number up to 100. Here’s how you fill out the Data Validation dialog to set up that data validation:

![Data Validation dialog](image)

**CAUTION**

Lists are case-sensitive. This is especially important where there isn’t an in-cell dropdown list. If you want to restrict entry in a cell to Yes or No, you have to decide whether other variations are acceptable, such as yes, YES, y, or Y and no, NO, n, or N.

If any or all of these are acceptable, you need to list them in the Source field in the Data Validation dialog. This figure shows the configuration for a list of acceptable entries in a cell that does not have an in-cell dropdown list:

![Data Validation dialog](image)
**TIP**
The Data Validation dialog has an Error Alert tab. As shown in the next figure, you can use this tab to advise users that they’ve tried to input an invalid entry and tell them what constitutes an acceptable entry.

![Data Validation Dialog](image)

In the figure below, data validation prevented the user from putting a 6 in cell C2 and showed your custom error message.

<table>
<thead>
<tr>
<th></th>
<th>Inspector</th>
<th>Date</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Preventing Start and End Times from Being Reversed

All the times in the following range are fine. But how do you prevent common errors with entering times? A test started at 9:00AM and ended at 1:00AM the same day is impossible...the user probably meant 1:00PM.
Without data validation, that 1:00AM would be accepted. To prevent this, you have to do two pieces of validation, as shown in the figures below:

- Prevent the start time from being larger than the end time.
- Prevent the end time from being smaller than the start time.

Data Validation Conclusions

Data validation is an invisible protective layer in a spreadsheet that helps keep people and spreadsheets out of trouble. When inaccurate data can create big problems, validation is one area where you cannot cut corners, especially when you are making a spreadsheet for other people to use. What’s most important about data validation is finding a balance between anticipating every possible error and understanding that you won’t be able to control everything.

Here are some common spreadsheet errors that you can anticipate:

- Addition of notes in a field that should have a Y or N
- Letters where only numbers should go (and vice versa)
- Inconsistent formatting and abbreviations
- Incorrect dates that are far in the future or past
- Incomplete email addresses
- Numbers that make no sense

Smart use of validation can eliminate or greatly reduce these problems and increase the trust you have in your results.
Implementing Error-Handling and Formula Triggers

Sometimes formulas give you weird results or you don’t know what certain errors mean. Rather than sit with a spreadsheet full of errors and gibberish, you can be proactive and build formulas that help manage results. Like data validation, error handling and formula triggers are critical to data analysts. Whether you use complex formulas or have painfully slow manual processes, you have to be able to trust the results.

Error-Handling Example

The following figure has a list in columns A:E. In this case, you want to retrieve the item names from the lookup range I3:J10. Column C uses a straight VLOOKUP, and two entries error out with #N/A.

Column E uses the IFERROR function to have Excel tell you when something is discontinued.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Item #</td>
<td>Quantity</td>
<td>Item Name</td>
<td>Item Name</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>CD10</td>
<td>20</td>
<td>Jeans</td>
<td>Jeans</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>VR92</td>
<td>9</td>
<td>Gloves</td>
<td>Gloves</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>VR339</td>
<td>24</td>
<td>#N/A</td>
<td>discontinued</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>CD23</td>
<td>11</td>
<td>Skirt</td>
<td>Skirt</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>KLD4</td>
<td>12</td>
<td>#N/A</td>
<td>discontinued</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>CD21</td>
<td>3</td>
<td>T-Shirt</td>
<td>T-Shirt</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>MM3K</td>
<td>17</td>
<td>Hat</td>
<td>Hat</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>CD23</td>
<td>5</td>
<td>Skirt</td>
<td>Skirt</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>CD21</td>
<td>31</td>
<td>T-Shirt</td>
<td>T-Shirt</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>MM1M</td>
<td>6</td>
<td>Sunglasses</td>
<td>Sunglasses</td>
<td></td>
</tr>
</tbody>
</table>

Formula Trigger Example

Say that you have a business rule which says that at least three readings are required for a final score. In the following figure, notice in column E that Test 2 has a final score with one reading, and Test 3 errors out because there are no readings (and you can’t divide by zero). In column F, you use IF and COUNTA as a trigger to determine whether there are three readings:

=IF(COUNTA(B2:D2)<3,””,AVERAGE((B2:D2))

Only if you have three readings will the average be taken.
Here’s what the formula above says to do:

- Count the non-empty cells in B2:D2.
- If the total is less than 3, don’t do anything.
- If the total is not less than 3, average the readings in B2:D2.

Error-Handling Functions: IFNA vs. IFERROR

IFNA allows you to work with formulas and errors—specifically #N/A errors—in a very granular way.

Let’s say you have a list of part numbers and need to use VLOOKUP to retrieve the part names from a master list. You also know that some part numbers don’t exist on the master list because they were discontinued. When VLOOKUP doesn’t find a part number, you don’t want the #N/A error; you’d like for the formula to return “discontinued.”

IFERROR is a huge help, but if you don’t use it carefully, it can mask too much. It masks #REF, #VALUE, circular references, among other things. Here’s an example:

```
=VLOOKUP(A5,$I$2:$K$7,w,FALSE)
```

This formula results in #NAME? because “w” makes no sense here; it was possibly a user typo. You need to see the #NAME? error because something is wrong with the formula. If you nested the formula in IFERROR, like this, it would hide the blunder and incorrectly tell you that all of your parts are discontinued:

```
=IFERROR(VLOOKUP(A5,$I$2:$K$7,w,FALSE),"discontinued")
```

However, wrapping the formula in IFNA, like this, gives you control over lookup values that are not found:

```
=IFNA(VLOOKUP(A5,$I$2:$K$7,3,FALSE),"discontinued")
```

**NOTE**

IFNA returns “discontinued” only if a part number isn’t found, and it still generates other error messages that warn you about the integrity of the formulas, data, or spreadsheet development.

**CAUTION**

While IFERROR works in Excel 2010 and later, the IFNA function was introduced in Excel 2013 and will not calculate in Excel 2010.
Using Pivot Charts

Pivot Charts are an amazing feature first introduced in Excel 2000.

**NOTE**
Sorry, Mac users. You don’t have access to Pivot Charts.

Pivot Charts have multiple benefits. Here are a few:

- They allow you to go straight to a chart that contains source data and skip over Pivot Tables. A Pivot Table is still created, so you get the interface that you’re used to seeing with Pivot Tables, but Excel builds a Pivot Chart at the same time.
- They provide interactivity. It’s like having multiple charts in the space of only one.
- A Pivot Chart provides a graph of your data without requiring you to summarize the data with a lot of formulas.

The following figure shows a spreadsheet that contains 100 rows of source data. The data is set in a table that has headers.

<table>
<thead>
<tr>
<th></th>
<th>Category</th>
<th>Origin</th>
<th>Transaction Amount</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Electronic Equipment</td>
<td>Germany</td>
<td>$1,004.40</td>
<td>European Union</td>
</tr>
<tr>
<td>2</td>
<td>Sporting Goods</td>
<td>USA</td>
<td>$651.50</td>
<td>North America</td>
</tr>
<tr>
<td>3</td>
<td>Electronic Equipment</td>
<td>South Korea</td>
<td>$3,192.00</td>
<td>Asia</td>
</tr>
<tr>
<td>4</td>
<td>Cutlery</td>
<td>USA</td>
<td>$2,489.22</td>
<td>North America</td>
</tr>
<tr>
<td>5</td>
<td>Food/Beverages</td>
<td>Pakistan</td>
<td>$53.82</td>
<td>Asia</td>
</tr>
<tr>
<td>6</td>
<td>Food/Beverages</td>
<td>Canada</td>
<td>$2,074.76</td>
<td>North America</td>
</tr>
<tr>
<td>7</td>
<td>Clothing</td>
<td>Brazil</td>
<td>$1,642.54</td>
<td>South America</td>
</tr>
</tbody>
</table>

It would be helpful to quickly create visuals from many different angles, such as the following:

- Transaction amount by region and category
- Count of transactions by category
- Count of transactions by certain categories and all origins

You can do this easily by making a Pivot Chart from the source data. You start by selecting Insert | Charts | Pivot Chart | Pivot Chart:
The Create PivotChart dialog, shown in the next figure, appears.

It defaults to the correct data range. Select the option New Worksheet so the Pivot Chart will be placed on a new worksheet, like this:
Look familiar? This looks like the Pivot Table interface, except that you have an empty chart in the middle of the page. You build Pivot Charts exactly the same way you build Pivot Tables: Drag items from the left side of the field list to the desired areas on the right side.

The following figure is a Pivot Chart that visually shows the amounts spent on each category, segmented by world region. Notice that the associated Pivot Table has been created behind the Pivot Chart. This can get annoying because every change you make to the Pivot Chart is also made to the Pivot Table.

You can avoid having the Pivot Chart on top of the Pivot Table by moving the Pivot Chart, like this:

1. Add a new worksheet.
2. Go back to the first worksheet and right-click the Pivot Chart Then select Move Chart.
3. In the Move Chart dialog that appears, choose Object In and use the dropdown list to choose your new worksheet. The following image shows that Chart1 is being moved to be an object in Sheet2.

![Move Chart dialog](image)

**Playing Around with a Pivot Chart**

The gray buttons on a Pivot Chart are filters, just like the ones you get on Pivot Tables. In the next figure, take a look at the categories Cutlery, Food/Beverage, and Sporting Goods. The Pivot Chart tells you that North American Cutlery has sold the greatest amount and Asian Sporting Goods has sold the second greatest amount.

![Pivot Chart](image)

To find out more detail about this, add Origin to the Axis section of the field list.
As you can see in the next figure, most of the money spent in the Asian Sporting Goods category went to South Korea.

![Graph showing transaction amounts by country and category]

**Adding New Data to a Pivot Chart**

To see what happens when you add new data to a Pivot Chart, add this new transaction to the Pivot Chart you've been working with:

- Food/Beverages
- Honduras
- $3,000
- Central America

This is what the Pivot Table looks like now, with this transaction added:

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>Sporting Goods</td>
<td>$</td>
<td>2,333.81 Europe Union</td>
</tr>
<tr>
<td>101</td>
<td>Clothing</td>
<td>$</td>
<td>867.00 South America</td>
</tr>
<tr>
<td>102</td>
<td>Food/Beverages</td>
<td>$</td>
<td>3,000.00 Central America</td>
</tr>
</tbody>
</table>

The Pivot Chart will not automatically reflect this new data. You have to refresh by right-clicking the Pivot Chart outside the plot area and selecting Refresh Data:
Here you can see that Honduras now shows up in the Central America section of the chart:

Changing the Chart Type

Just as you can customize regular charts, you can customize Pivot Charts. To see how this works, in this section, you’ll change the default clustered column charts to stacked column charts so that you can see overall totals for all categories. Follow these steps:

1. Remove Origin from the Axis section of the field list.
2. Clear the Category filter by clicking the button and choosing Select All.
3. Click on the Pivot Chart so that the Pivot Chart Tools tab appears in the ribbon:
4. Select Design | Change Chart Type.
5. In the Change Chart Type dialog, select stacked column, and then click OK.
Using Slicers with Pivot Charts

Dealing with the field buttons on a chart can be a clumsy process. You can get rid of them and use slicers instead. Here’s how you add slicers:

1. Right-click any one of the field buttons on the Pivot Chart and select Hide All Field Buttons on Chart.
2. Click anywhere inside the Pivot Chart.
3. Select Insert | Filters | Slicer.
4. In the Slicer dialog, select the Category, Origin, and Region check boxes.

Excel creates these three slicers for you. Go ahead and rearrange, resize and configure them to suit your needs. For example, you can use the slicers to show just:

- Region: Asia and Central America
- Category: Art Supplies, Clothing, and Cutlery

This is what you get:

Now that I’ve whetted your appetite for slicers, they are the focus of the next section.
Using Slicers

Think of slicers, which became available with Excel 2010, as filters that easily allow you to slice into data without having to deal with dropdown filters. Here you have a range of data about 100 employees in a table:

<table>
<thead>
<tr>
<th>A</th>
<th>Training Status</th>
<th>B</th>
<th>Started</th>
<th>C</th>
<th>Grade</th>
<th>D</th>
<th>Phone Ext</th>
<th>E</th>
<th>Supervisor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Name</td>
<td>2</td>
<td>Adam Kenney</td>
<td>Failed</td>
<td>August</td>
<td>58</td>
<td>6703 Marcia Lovell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Walter Cox</td>
<td>4</td>
<td>Roy &quot;Scooter&quot; Green</td>
<td>Not Required</td>
<td>NR</td>
<td>5440 Abe Garcia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Jake Washington</td>
<td>6</td>
<td>Jack Evans Riley</td>
<td>Failed</td>
<td>August</td>
<td>65</td>
<td>3609 Abe Garcia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Kimberly Murphy</td>
<td>8</td>
<td>Lawrence Nelson</td>
<td>Not Required</td>
<td>NR</td>
<td>5485 Marcia Lovell</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Daniel Long</td>
<td>10</td>
<td>Evelyn James, CPA</td>
<td>Retake Scheduled</td>
<td>October</td>
<td>2746 Abe Garcia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Henry Alexander</td>
<td>12</td>
<td>Gloria Russell</td>
<td>Passed</td>
<td>October</td>
<td>78</td>
<td>5079 Marcia Lovell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Barbara Mitchell</td>
<td>14</td>
<td>Margaret S. Garcia</td>
<td>Passed</td>
<td>August</td>
<td>87</td>
<td>5714 Abe Garcia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Timothy Morris</td>
<td>16</td>
<td>David Collins</td>
<td>Failed</td>
<td>August</td>
<td>65</td>
<td>1354 János Crosz</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

You can create a Pivot Table to look closely at the activity around this training. (See page 79 for information on creating a Pivot Table.) Use the following information to create the Pivot Table shown below:

- Columns: Training Status
- Rows: Name
- Values: Average of Grade (See page 87 for changing to Average instead of Count or Sum.)
Now select Insert | Filters | Slicers. The Insert Slicers dialog appears. From it, choose the following (see the next figure):

- Training Status
- Started
- Supervisor

Notice the Slicer tab that appears on the ribbon only when a slicer is selected. You can use this tab to change the slicer colors, align slicers, and create multiple columns:

Here is your brand-new interface, where you can slice into the data as needed:

You can see in this figure that the slicers are filtering to show employees who fit the following criteria:

- Supervised by Gail Nurmi
- Training Status is either Passed or Failed
NOTE
A red x next to a funnel indicates that a slicer is active. On the Started slicer, there is no red x, but two choices are grayed out: Blank and NR. They signify that there is no data for those choices, based on what’s been filtered out.

NOTE
To select more than one item in a slicer, hold down the Ctrl key as you make your choices. To select adjacent items, click on the first and drag to the last.

Excel 2013: Guerrilla Data Analysis Gets Real

There are lots of new features in Excel 2013! Some people groan and others cheer—and many of them for the same reason: They wonder if Excel is trying to turn into a database. Possibly. Excel developers seem to understand that people need database functionality, but for any number of reasons, users continue to find ways to use Excel as a database that’s good enough to suit their purposes.

Let’s look at two features of Excel 2013 that are almost in database-land.

Using Slicers with Tables in Excel 2013

You’ve already seen how to use slicers with Pivot Tables, and up until Excel 2013, that was the only way to use them. But sometimes you don’t need a Pivot Table. Say that you have the data shown in the following figure, but you just want the phone extensions of the people whose Training Status is Passed and who started in October. You shouldn’t need a whole Pivot Table to dig out that much information.
The following figure shows a slice applied to find that information. Notice that the total row is toggled on, the count in column F is 14, and the average grade is 82.428.
Understanding Data Models and Relationships

You’ve seen how you can merge data from multiple data sources. Excel 2013 provides a fantastic tool for leaving data in place while treating the separate ranges as one range: data models.

For example, the following figure shows a worksheet whose data includes:

- Sales
- Sales representatives and their assistants
- Sales regions and the managers over those regions

Your goal is to pay a bonus to the assistants with the most sales.

You can spend time merging these three sources of data, but why bother? You can instead create relationships and make your work easy.

First, notice the names of the tables. With the cursor anywhere within a table, you can look at the Table Design tab in the Properties group to see the table’s name. The following figure shows that with the cursor in F2, the Properties group shows that you’re in Table1.

Here you’re working with two other tables:

- Table2: Region
- Table3: Sales Reps & Assistants

**TIP**
Excel’s default names for tables are Table1, Table2, etc. You can change table names to make formulas easier to understand just by typing over the default
table name and adding a brief, descriptive title. In this image, the table has been renamed Reps

Here’s how you create relationships between these tables:

2. In the Create Relationship dialog, create the first relationship by linking Territory in Table1 to Region in Table2 (see the next figure). Click OK.
3. Repeat step 1 and then, in the Create Relationship dialog, create another new relationship that links Table1 Rep and Table 3 Sales Rep. Click OK. You now have the relationships shown in the next figure.

Next, you’ll create a data model and tally the total sales for the sales assistants. Here’s what you do:

1. Select Insert | Pivot Table.
2. In the Create PivotTable dialog that appears (see the next figure), check Add This Data to the Data Model box and click OK.
3. From here you use the interface the same way you use the Pivot Table interface. Choose Amount in Table1 and Assistant in Table3. The Pivot Table shows you that Fred generated more sales than Aaron:

You can take this one step further by rearranging the data to see how the sales break down by assistant and manager (see the next figure). Now you’re using information from all three tables, and you haven’t used formulas or rearranged the source data.