LEARNING MADE EASY

5th Edition

Microsoft® Excel® Data Analysis

for dummies®

A Wiley Brand

Master Excel's powerful data-analysis tools

Generate data insights with PivotTables and trend forecasts

Visualize complex data with graphics and PivotCharts

Paul McFedries

Bestselling author of more than 100 books

Excel® Data Analysis

5th Edition

by Paul McFedries

for dummies®
A Wiley Brand

Excel® Data Analysis For Dummies®, 5th Edition

Published by: **John Wiley & Sons, Inc.,** 111 River Street, Hoboken, NJ 07030-5774, www.wiley.com

Published simultaneously in Canada

For general information on our other products and services, please contact our Customer Care Department within the U.S. at 877-762-2974, outside the U.S. at 317-

572-3993, or fax 317-572-4002. For technical support, please visit https://hub.wiley.com/community/support/dummies.

Wiley publishes in a variety of print and electronic formats and by print-on-demand. Some material included with standard print versions of this book may not be included in e-books or in print-on-demand. If this book refers to media such as a CD or DVD that is not included in the version you purchased, you may download this material at http://booksupport.wiley.com. For more information about Wiley products, visit www.wiley.com.

Library of Congress Control Number is available from the publisher.

ISBN 978-1-119-84442-6 (pbk); ISBN 978-1-119-84446-4 (ebk); ISBN 978-1-119-84447-1 (ebk)

Excel® Data Analysis For Dummies®

To view this book's Cheat Sheet, simply go to www.dummies.com and search for “Excel Data Analysis For Dummies Cheat Sheet” in the Search box.

Table of Contents

List of Tables

Chapter 3

Chapter 12

List of Illustrations

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

Chapter 9

Chapter 10

Chapter 11

Chapter 12

Chapter 13

Chapter 14

Chapter 16

Introduction

The world is bursting at the seams with data. It's on our computers, it's in our networks, it's on the web. Some days, it seems to be in the very air itself, borne on the wind. But here's the thing: No one actually cares about data. A collection of data — whether it resides on your PC or some giant server somewhere — is really just a bunch of numbers and text, dates and times. No one cares about data because data doesn't *mean* anything. Data isn't cool. You know what's cool? *Knowledge* is cool. *Insight* is cool.

So how do you turn data into knowledge? How do you tweak data to generate insight? You need to organize that data, and then you need to sort it, filter it, run calculations on it, and summarize it. In a word, you need to *analyze* the data.

Now for the good news: If you have (or can get) that data into Excel, you have a giant basket of data-analysis tools at your disposal. Excel really seems to have been made with data analysis in mind, because it offers such a wide variety of features and techniques for organizing, manipulating, and summarizing just about anything that resides in a worksheet. If you can get your data into Excel, it will help you turn that data into knowledge and insight.

This book takes you on a tour of Excel's data-analysis tools. You learn everything you need to know to make your data spill its secrets and to uncover your data's hidden-in-plain-sight wisdom. Best of all, if you already know how to perform the basic Excel chores, you don't need to learn any other fancy-schmancy Excel techniques to get started in data analysis. Sweet? You bet.

About This Book

This book contains 16 chapters (and a bonus appendix), but that doesn't mean that you have to, as the King says gravely in *Alice's Adventures in Wonderland,* "Begin at the beginning and go on till you come to the end: Then stop." If you've done a bit of data-analysis work in the past, please feel free to dip into the book wherever it strikes your fancy. The chapters all present their data-analysis info and techniques in readily digestible, bite-sized chunks, so you can certainly graze your way through this book.

However, if you're brand spanking new to data analysis — particularly if you're not even sure what data analysis even *is* — no problem: I'm here to help. To get your data-analysis education off to a solid start, I highly recommend reading the book's first three chapters to get some of the basics down cold. From there, you can travel to more advanced territory, safe in the knowledge that you've got some survival skills to fall back on.

What You Can Safely Ignore

This book consists of several hundred pages. Do I expect you to read every word on every page? Yes, I do. Just kidding! No, of course I don't. Entire sections — heck, maybe even entire *chapters* — might contain information that's not relevant to what you do. That's fine and my feelings won't be hurt if you skim through (or — who's kidding whom? — skip over) those parts of the book.

If time (or attention) is short, what else might you want to ignore? Okay, in many places throughout the book I

provide step-by-step instructions to complete some task. Each of those steps includes some bold type that gives you the basic instruction. In many cases, however, below that bold text I offer supplementary information to flesh out or extend or explain the bold instruction. Am I just showing off how much I know about all this stuff? Yes, sometimes. Do you have to read these extended instructions? Nope. Read the bold stuff, for sure, but feel free to skip the details if they seem unnecessary or unimportant.

Foolish Assumptions

This book is for people who are new (or relatively new) to Excel data analysis. That doesn't mean, however, that the book is suitable to people who have never used a PC, Microsoft Windows, or even Excel. So first I assume not only that you have a PC running Microsoft Windows but also that you've had some experience with both. (For the purposes of this book, that just means you know how to start and switch between programs.) I also assume that your PC has a recent version of Excel installed. What does "recent" mean? Well, this book is based on Excel 2021, but you should be fine if you're running Excel 365, Excel 2019, Excel 2016, or even Excel 2013.

As I said before, I do *not* assume that you're an Excel expert, but I do assume that you know at least the following Excel basics:

- Creating, saving, opening, and switching between workbooks
- Creating and switching between worksheets
- Finding and running commands on the Ribbon

- Entering numbers, text, dates, times, and formulas into worksheet cells
- Working with Excel's basic worksheet functions

Icons Used in This Book

Like other books in the *For Dummies* series, this book uses icons, or little margin pictures, to flag things that don't quite fit into the flow of the chapter discussion. Here are the icons that I use:

REMEMBER This icon marks text that contains some things that are useful or important enough that you'd do well to store the text somewhere safe in your memory for later recall.

TECHNICAL STUFF This icon marks text that contains some for-nerds-only technical details or explanations that you're free to skip.

TIP This icon marks text that contains a shortcut or an easier way to do things, which I hope will make your life — or, at least, the data-analysis portion of your life — more efficient.

WARNING This icon marks text that contains a friendly but unusually insistent reminder to avoid doing something. You have been warned.

Beyond the Book

- **Examples:** This book's sample Excel workbooks can be found by going to www.dummies.com/go/exceldataanalysisfd5e or at my website: www.paulmcfedries.com.
- **Cheat Sheet:** To locate this book's cheat sheet, go to www.dummies.com and search for *Excel Data Analysis For Dummies.* See the cheat sheet for info on Excel database functions, Boolean expressions, and important statistical terms.
- **Updates:** If this book has any updates after printing, they will be posted to this book's page at www.dummies.com.

Where to Go from Here

If you're just getting your feet wet with Excel data analysis, flip the page and start perusing the first chapter.

If you have some experience with Excel data analysis or you have a special problem or question, use the Table of Contents or the index to find out where I cover that topic and then turn to that page.

Either way, happy analyzing!

Part 1

Getting Started with Data Analysis

IN THIS PART ...

Understand data analysis and get to know basic analysis features such as conditional formatting and subtotals.

Discover Excel’s built-in tools for analyzing data.

Learn how to build Excel tables that hold and store the data you need to analyze.

Find quick and easy ways to begin your analysis using simple statistics, sorting, and filtering.

Get practical stratagems and common-sense tactics for grabbing data from extra sources.

Chapter 1

Learning Basic Data-Analysis Techniques

IN THIS CHAPTER

- **Learning about data analysis**
- **Analyzing data by applying conditional formatting**
- **Adding subtotals to summarize data**
- **Grouping related data**
- **Combining data from multiple worksheets**

You are awash in data. Information multiplies around you so fast that you wonder how to make sense of it all. You think, “I know what to do. I'll paste the data into Excel. That way, at least the data will be nicely arranged in the worksheet cells, and I can add a little formatting to make things somewhat palatable.” That’s a fine start, but you’re often called upon to do more with your data than make it merely presentable. Your boss, your customer, or perhaps just your curiosity requires you to divine some inner meaning from the jumble of numbers and text that litter your workbooks. In other words, you need to *analyze* your data to see what nuggets of understanding you can unearth.

This chapter gets you started down that data-analysis path by exploring a few straightforward but useful analytic techniques. After discovering what data analysis entails, you investigate a number of Excel data-analysis

techniques, including conditional formatting, data bars, color scales, and icon sets. From there, you dive into some useful methods for summarizing your data, including subtotals, grouping, and consolidation. Before you know it, that untamed wilderness of a worksheet will be nicely groomed and landscaped.

What Is Data Analysis, Anyway?

Are you wondering, “What is data analysis, anyway?” That’s an excellent question! Here’s an answer that I unpack for you as I go along: *Data analysis* is the application of tools and techniques to organize, study, reach conclusions, and sometimes make predictions about a specific collection of information.

For example, a sales manager might use data analysis to study the sales history of a product, determine the overall trend, and produce a forecast of future sales. A scientist might use data analysis to study experimental findings and determine the statistical significance of the results. A family might use data analysis to find the maximum mortgage it can afford or how much it must put aside each month to finance retirement or the kids’ education.

Cooking raw data

The point of data analysis is to understand information on some deeper, more meaningful level. By definition, *raw data* is a mere collection of facts that by themselves tell you little or nothing of any importance. To gain some understanding of the data, you must manipulate the data in some meaningful way. The purpose of manipulating data can be something as simple as finding the sum or

average of a column of numbers or as complex as employing a full-scale regression analysis to determine the underlying trend of a range of values. Both are examples of data analysis, and Excel offers a number of tools — from the straightforward to the sophisticated — to meet even the most demanding needs.

Dealing with data

The *data* part of *data analysis* is a collection of numbers, dates, and text that represents the raw information you have to work with. In Excel, this data resides inside a worksheet, which makes the data available for you to apply Excel's satisfyingly large array of data-analysis tools.

Most data-analysis projects involve large amounts of data, and the fastest and most accurate way to get that data onto a worksheet is to import it from a non-Excel data source. In the simplest scenario, you can copy the data from a text file, a Word table, or an Access datasheet and then paste it into a worksheet. However, most business and scientific data is stored in large databases, so Excel offers tools to import the data you need into your worksheet. I talk about all this in more detail later in the book.

After you have your data in the worksheet, you can use the data as is to apply many data-analysis techniques. However, if you convert the range into a *table,* Excel treats the data as a simple database and enables you to apply a number of database-specific analysis techniques to the table.

Building data models

In many cases, you perform data analysis on worksheet values by organizing those values into a *data model,* a collection of cells designed as a worksheet version of

some real-world concept or scenario. The model includes not only the raw data but also one or more cells that represent some analysis of the data. For example, a mortgage amortization model would have the mortgage data — interest rate, principal, and term — and cells that calculate the payment, principal, and interest over the term. For such calculations, you use formulas and Excel's built-in worksheet functions.

Performing what-if analysis

One of the most common data-analysis techniques is *what-if analysis,* for which you set up worksheet models to analyze hypothetical situations. The "what-if" part means that these situations usually come in the form of a question: "What happens to the monthly payment if the interest rate goes up by 2 percent?" "What will the sales be if you increase the advertising budget by 10 percent?" Excel offers four what-if analysis tools: data tables, Goal Seek, Solver, and scenarios, all of which I cover in this book.

Analyzing Data with Conditional Formatting

Many Excel worksheets contain hundreds of data values. You could try to make sense of such largish sets of data by creating complex formulas and wielding Excel's powerful data-analysis tools. However, just as you wouldn't use a steamroller to crush a tin can, sometimes these sophisticated techniques are too much tool for the job. For example, what if all you want are answers to simple questions such as the following:

- Which cell values are less than 0?