The Excel® Analyst's Guide to Access®

Michael Alexander

Table of Contents

Title Page

Copyright

Dedication

About the Author

Credits

Acknowledgments

Introduction

- *What to Expect from This Book*
- *What Not to Expect from This Book*
- *Skills Required for This Book*
- *How This Book is Organized*
- *Conventions*
- *Companion Database*

Part I: Fundamentals of Data Analysis in Access

Chapter 1: The Case for Data Analysis in Access

Where Data Analysis with Excel Can Go Wrong
Deciding Whether to Use Access or Excel
An Excel User's Guide to Access: Don't Panic!
Summary

Chapter 2: Access Basics

Access Tables
Understanding the Relational Database Concept
Query Basics
The Top Ten Query Errors
Summary

Chapter 3: Beyond Select Queries

Aggregate Queries
Action Queries
Crosstab Queries
Summary

Part II: Basic Analysis Techniques

Chapter 4: Transforming Your Data with Access

Finding and Removing Duplicate Records
Common Transformation Tasks

Summary

Chapter 5: Working with Calculations and Dates

Using Calculations in Your Analysis
Using Dates in Your Analysis
The New Calculated Data Type
Summary

Chapter 6: Performing Conditional Analysis

Using Parameter Queries
Using Conditional Functions
Summary

Part III: Advanced Analysis Techniques

Chapter 7: Adding Dimension with Subqueries and Domain Aggregate Functions

Enhancing Your Analysis with Subqueries
Domain Aggregate Functions
Summary

Chapter 8: Running Descriptive Statistics in Access

Basic Descriptive Statistics

Advanced Descriptive Statistics
Summary

Chapter 9: Scheduling and Running Batch Analysis

Introduction to Access Macros
Setting Up and Managing Batch Analysis
Scheduling Macros to Run Nightly
Summary

Chapter 10: Leveraging VBA to Enhance Data Analysis

Creating and Using Custom Functions
Controlling Analytical Processes with Forms
Processing Data Behind the Scenes
Summary

Part IV: Reports, Dashboards, and Visualization in Access

Chapter 11: Presenting Data with Access Reports

Access Report Basics
Creating and Modifying Grouped Reports
Creating a Report from Scratch
Summary

Chapter 12: Using Pivot Tables and Pivot Charts in Access

- Pivot Tables in Access?
- The Anatomy of a Pivot Table
- Creating a Basic Pivot Table
- Creating an Advanced Pivot Table with Details
- Saving Your Pivot Table
- Sending Your Access Pivot Table to Excel
- Pivot Table Options
- Working with Pivot Charts in Access
- Summary

Chapter 13: Enhancing Queries and Reports with Visualizations

- Basic Visualization Techniques
- Advanced Visualization Techniques
- Summary

Part V: Advanced Excel and Access Integration Techniques

Chapter 14: Getting Access Data into Excel

- The Different Methods for Importing Access Data
- Summary

Chapter 15: Using VBA to Move Data between Excel and Access

- Understanding ADO Fundamentals
- Common Scenarios Where VBA Can Help
- Summary

Chapter 16: Exploring Excel and Access Automation

- Understanding the Concept of Binding
- Automating Excel from Access
- Automating Access from Excel
- Summary

Chapter 17: Integrating Excel and Access with XML

- Why XML?
- Understanding XML
- Creating a Simple Reporting Solution with XML
- Creating a Data Entry Process Using XML
- Summary

Chapter 18: Integrating Excel and Other Office Applications

- Integrating Excel with Microsoft Word
- Integrating Excel with PowerPoint
- Integrating Excel and Outlook
- Summary

Part VI: Appendixes

Appendix A: Access VBA Fundamentals

- Covering the Basics in 10 Steps
- Letting Access Teach You VBA

Appendix B: Understanding and Using SQL

- Understanding Basic SQL
- Getting Fancy with Advanced SQL Statements
- Using SQL Specific Queries

Appendix C: Query Performance, Database Corruption, and Other Thoughts

- Optimizing Query Performance
- Handling Database Corruption
- Getting Help in Access

Appendix D: Data Analyst's Function Reference

- Abs
- Asc
- Atn
- Choose
- Chr

Cos
Date
DateAdd
DateDiff
DatePart
DateSerial
DateValue
Day
DDB
Domain Aggregate Functions
Exp
FormatCurrency
FormatDateTime
FormatNumber
FormatPercent
FV
Hour
IIf
InStr
InStrRev
IPmt
IRR
IsError
IsNull
IsNumeric
LCase
Left
Len
Log

Mid
Minute
MIRR
Month
MonthName
Now
NPer
NPV
NZ
Partition
Pmt
PPmt
PV
Rate
Replace
Right
Rnd
Round
Second
Sgn
Sin
SLN
Space
SQL Aggregate Functions
Sqr
Str
StrConv
String
StrReverse

Switch
SYD
Tan
Time
TimeSerial
TimeValue
Trim, LTrim, RTrim
TypeName
UCase
Val
VarType
Weekday
WeekdayName
Year

Index

The Excel® Analyst's Guide to Access®

Michael Alexander

WILEY
Wiley Publishing, Inc.

The Excel® Analyst's Guide to Access®

Published by

Wiley Publishing, Inc.

10475 Crosspoint Boulevard

Indianapolis, IN 46256

www.wiley.com

Copyright © 2010 by Wiley Publishing, Inc., Indianapolis, Indiana

Published simultaneously in Canada

ISBN: 978-0-470-56701-2

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at http://www.wiley.com/go/permissions.

Limit of Liability/Disclaimer of Warranty: The publisher and the author make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation warranties of fitness for a particular purpose. No warranty may be created or extended by sales or promotional materials. The advice and strategies contained herein may not be

suitable for every situation. This work is sold with the understanding that the publisher is not engaged in rendering legal, accounting, or other professional services. If professional assistance is required, the services of a competent professional person should be sought. Neither the publisher nor the author shall be liable for damages arising herefrom. The fact that an organization or Web site is referred to in this work as a citation and/or a potential source of further information does not mean that the author or the publisher endorses the information the organization or Web site may provide or recommendations it may make. Further, readers should be aware that Internet Web sites listed in this work may have changed or disappeared between when this work was written and when it is read.

For general information on our other products and services please contact our Customer Care Department within the United States at (877) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Control Number: 2010922044

Trademarks: Wiley and the Wiley logo are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates, in the United States and other countries, and may not be used without written permission. Excel and Access are registered trademarks of Microsoft Corporation in the United States and/or other countries. All other trademarks are the property of their respective owners. Wiley Publishing, Inc. is not associated with any product or vendor mentioned in this book.

This is dedicated to the fans of DataPigTechnologies.com...all 12 of you.

About the Author

Mike Alexander is a Microsoft Certified Application Developer (MCAD) and author of several books on advanced business analysis with Microsoft Access and Excel. He has more than 15 years experience consulting and developing Office solutions. Michael has been named a Microsoft MVP for his ongoing contributions to the Excel community.

In his spare time he runs a free tutorial site, www.datapigtechnologies.com, where he shares basic Access and Excel tips to the Office community.

Credits

Executive Editor
Carol Long

Project Editor
Maureen Spears

Technical Editor
Dick Kusleika

Production Editor
Kathleen Wisor

Copy Editor
C.M. Jones

Editorial Director
Robyn B. Siesky

Editorial Manager
Mary Beth Wakefield

Marketing Manager
Ashley Zurcher

Production Manager
Tim Tate

Vice President and Executive Group Publisher
Richard Swadley

Vice President and Executive Publisher
Barry Pruett

Associate Publisher
Jim Minatel

Project Coordinator, Cover
Lynsey Stanford

Proofreader
Sheilah Ledwidge, Word One

Indexer
Robert Swanson

Cover Designer
Ryan Sneed

Acknowledgments

My deepest thanks to Carol Long and Maureen Spears, for all the hours of work put into bringing this book to life. Thanks also to Dick Kusleika (one of the best tech editors I've worked with) for suggesting numerous improvements to the examples and text in this book. Finally, a special thank you goes out to the wife and kids for putting up with all the time spent locked away on this project.

Introduction

If you were to ask a random sampling of people what data analysis is, most would say that it is the process of calculating and summarizing data to get an answer to a question. In one sense, they are correct. However, the actions they are describing represent only a small part of the process known as data analysis.

For example, if you were asked to analyze how much revenue in sales your company made last month, what would you have to do in order to complete that analysis? You would just calculate and summarize the sales for the month, right? Well, where would you get the sales data? Where would you store the data? Would you have to clean up the data when you got it? How would you present your analysis: by week, by day, by location? The point being made here is that the process of data analysis is made up of more than just calculating and summarizing data.

A more representative definition of data analysis is the process of systematically collecting, transforming, and analyzing data in order to present meaningful conclusions. To better understand this concept, think of data analysis as a process that encapsulates four fundamental actions: collection, transformation, analysis, and presentation.

- **Collection.** Collection encompasses the gathering and storing of data—that is, where you obtain your data, how you will receive your data, how you will store your data, and how you will access your data when it comes time to perform some analysis.
- **Transformation.** Transformation is the process of ensuring your data is uniform in structure, free from

redundancy, and stable. This generally entails things like establishing a table structure, cleaning text, removing blanks, and standardizing data fields.

- **Analysis.** Analysis is the investigation of the component parts of your data and their relationships to your data source as a whole. You are analyzing your data when you are calculating, summarizing, categorizing, comparing, contrasting, examining, or testing your data.
- **Presentation.** In the context of data analysis, presentation deals with how you make the content of your analysis available to a certain audience. That is, how you choose to display your results. Some considerations that go along with presentation of your analysis include the platform you will use, the levels of visibility you will provide, and the freedom you will give your audience to change their view.

As you think about these four fundamental actions, think about this reality: most analysts are severely limited to one tool—Excel. This means that all of the complex actions involved in each of these fundamentals are mostly being done with and in Excel. What's the problem with that? Well, Excel is not designed to do many of these actions. However, many analysts are so limited in their toolsets that they often go into hand-to-hand combat with their data, creating complex workarounds and inefficient processes.

What this book highlights is that there are powerful functionalities in Access that can help you go beyond your one dimensional spreadsheet and liberate you from the daily grind of managing and maintaining redundant analytical processes. Indeed, using Access for your data analysis needs can help you streamline your analytical processes, increase your productivity, and analyze the larger datasets that have reached Excel's limitations.

Throughout this book, you will come to realize that Access is not the dry database program used only for storing data and building departmental applications. Access possesses strong data analysis functionalities that are easy to learn and certainly applicable to many types of organizations and data systems.

What to Expect from This Book

After reading the first three chapters, you will be able to demonstrate proficiency in Access, executing powerful analysis on large datasets that have long since reached Excel's limitations. After the first nine chapters, you'll be able to add depth and dimension to your analysis with advanced Access functions, building complex analytical processes with ease. By the end of the book, you'll be creating your own custom functions, performing batch analysis, and developing automated procedures that essentially run on their own. You'll also you will be able to analyze large amounts of data in a meaningful way, quickly slice data into various views on the fly, automate redundant analysis, save time, and increase productivity.

What Not to Expect from This Book

It's important to note that there are aspects of Access and data analysis that are out of the scope of this book.

While this book does cover the fundamentals of Access, it is always in the light of data analysis and it is written from a data analyst's point of view. This is not meant to be an all-encompassing book on Access. That being said, if you are a first-time user of Access, you can feel confident that this book will provide you with a solid

introduction to Access that will leave you with valuable skills you can use in your daily operations.

This book is not meant to be a book on data management theory and best practices. Nor is it meant to expound on high-level business intelligence concepts. This is more of a “technician's” book, providing hands-on instruction that introduces Access as an analytical tool that can provide powerful solutions to common analytical scenarios and issues.

Finally, while this book does contain a chapter that demonstrates various techniques to perform a whole range of statistical analysis, it is important to note that this book does not cover statistics theory, methodology, or best practices.

Skills Required for This Book

In order to get the most out of this book, it's best that you have certain skills before diving into the topics highlighted in this book. The ideal candidate for this book will have:

- Some experience working with data and familiarity with the basic concepts of data analysis such as working with tables, aggregating data, and performing calculations
- Experience using Excel with a strong grasp of concepts such as table structures, filtering, sorting and using formulas
- Some basic knowledge of Access; enough to know it exists and to have opened a database once or twice

How This Book is Organized

The following sections discuss this books structure and what it has to offer.

Part I: Fundamentals of Data Analysis in Access

Part I, which includes Chapters 1, 2 and 3, provides a condensed introduction to Access. Here, you will learn some of the basic fundamentals of Access, along with the essential query skills required throughout the rest of the book. Topics covered in this part are: relational database concepts, query basics, using aggregate queries, action queries, and Crosstab queries.

Part II: Basic Analysis Techniques

Part II introduces you to some of the basic analytical tools and techniques available in Access. Chapter 4 covers data transformation, providing examples of how to clean and shape raw data into staging areas. Chapter 5 provides in-depth instruction on how to create and utilize custom calculations in analysis. Chapter 5 also shows you how to work with dates, using them in simple date calculations. Chapter 6 introduces you to some conditional analysis techniques that allow for the addition of business logic into analytical processes.

Part III: Advanced Analysis Techniques

Part III demonstrates many of the advanced techniques that truly bring data analysis to the next level. Chapter 7 introduces you to powerful subquery and domain aggregate functionality. Chapter 8 demonstrates many of the advanced statistical analysis that can be performed using subqueries and domain aggregate functions.

Chapter 9 provides you with an in-depth look at Access macros and how to schedule batch data processing. Chapter 10 not only shows you how to use SQL and VBA to run data analysis without queries, but also how to create your own custom functions.

Part IV: Reports, Dashboards and Visualizations in Access

Part IV focuses on building reports and visualizations using Access. In Chapter 11, you will cover the basics of turning data into a slick-looking PDF-style Access reports. The chapter also talks about creating charts in Access to enhance the look and feel of Access reports. Chapter 12 discusses the real-world benefits of using the built-in PivotTable and PivotChart functionality found in Access. Chapter 13 demonstrates some of the innovative ways you can implement dashboard-style visualizations in your Access Queries and Reports.

Part V: Advanced Excel and Access Integration Techniques

Part V turns your attention to automation and integration, showing you how your reporting mechanisms can be enhanced by leveraging other programs and platforms. Chapter 14 discusses the various ways to move data between Excel and Access using VBA and ADO. Chapter 15 focuses on using Excel and Access automation to manage the inevitable need to show parts of your reporting through Excel. In Chapter 16, you're introduced to the automation techniques, which allow Excel and Access to take control of one another, resulting in some interesting reporting options. In Chapter 17, you get a thorough introduction to XML including a detailed explanation of how XML can collect and transfer data. You

conclude with Chapter 18, where you'll get a taste of some of the techniques you can use to integrate Excel and other applications in the Microsoft Office suite. Here, you will be shown how to perform the most common tasks in more efficient ways through integration.

Part VI: Appendixes

Part VI includes useful reference materials that will assist you in your everyday dealings with Access. Appendix A provides a high-level overview of VBA for those users who are new to the world of Access programming. Appendix B introduces SQL, offering a concise tutorial on SQL syntax and usage. Appendix C offers ideas on how to avoid performance and corruption issues when working with Access databases. Appendix D details many of the built-in Access functions that are available to data analysts.

Conventions

To help you get the most from the text and keep track of what's happening, we've used a number of conventions throughout the book.

TRICKS OF THE TRADE

These features give you in-depth information on how to handle specific problems you might encounter when working with Access.

WARNING

These hold important, not-to-be forgotten information that is directly relevant to the surrounding text.

> **NOTE**
>
> **Notes, tips, hints, tricks, and asides to the current discussion are presented like this.**

As for styles in the text:

We *highlight* new terms and important words when we introduce them.

- We show keyboard strokes like this: Ctrl+A.
- We show file names, URLs, and code within the text like so: `persistence.properties`.
- We present code as follows:

```
We use a monofont type for code examples.
```

Companion Database

The examples demonstrated throughout this book can be found in the companion database. This sample database is located at www.wiley.com/go/excelanalystguide.

Part I

Fundamentals of Data Analysis in Access

In this Part

Chapter 1: The Case for Data Analysis in Access
Chapter 2: Access Basics
Chapter 3: Beyond Select Queries

Chapter 1

The Case for Data Analysis in Access

When you ask most people which software tool they use for their daily data analysis, the answer you most often get is Excel. Indeed, if you were to enter the key words *data analysis* in an Amazon.com search, you would get a plethora of books on how to analyze your data with Excel. Well if so many people seem to agree that using Excel to analyze data is the way to go, why bother using Access for data analysis? The honest answer: to avoid the limitations and issues that plague Excel.

This is not meant to disparage Excel. Many people across varying industries have used Excel for years, considering it the tool of choice for performing and presenting data analysis. Anyone who does not understand Excel in today's business world is undoubtedly hiding that shameful fact. The interactive, impromptu analysis that Excel can perform makes it truly unique in the industry.

However, Excel is not without its limitations, as you will see in the following section.

Where Data Analysis with Excel Can Go Wrong

Years of consulting experience have brought me face to face with managers, accountants, and analysts who all