

FORENSIC ANALYTICS

Methods and Techniques for
Forensic Accounting Investigations

MARK J. NIGRINI, PHD

Forensic Analytics

*Methods and Techniques
for Forensic Accounting
Investigations*

MARK J. NIGRINI,
B.COM.(HONS), MBA, PH.D.



WILEY

John Wiley & Sons, Inc.

Preface

THE BUSINESS OF OCCUPATIONAL and financial statement fraud is unfortunately alive and doing very well. There are regular reports of financial statement fraud in the financial press, and all types of financial fraud in the press releases section of the SEC's website. There are also regular reports of occupational fraud in the financial press. These reports might just be the tip of the iceberg. The 2010 *Report to the Nations on Occupational Fraud and Abuse* of the Association of Certified Fraud Examiners estimates that the typical organization loses 5 percent of its annual revenue to fraud. These statistics are confirmed in other fraud surveys such as *The Global Economic Crime Survey* of PriceWaterhouseCoopers (2009) and in reports published by the U.S. Government Accountability Office. Together with the losses from employee fraud, there are also other corporate and public sector losses from accounting errors such as underbilling or overpaying or duplicate payments.

Forensic analytics describes the act of obtaining and analyzing electronic data using formulas and statistical techniques to reconstruct, detect, or otherwise support a claim of financial fraud. In this book, forensic analytics is also used to detect accounting errors such as underbilling or overpayments. Forensic analytics also includes the detection of biases that come about when people aim for specific numbers or number ranges to circumvent actual or perceived internal control thresholds. The use of forensic analytics has been made easier with the continued increase in computing power available on laptop computers and access to inexpensive software capable of some rigorous data analysis on large data tables. The main steps in forensic analytics are (a) data collection, (b) data preparation, (c) the use of forensic analytics, and (d) evaluation, investigation, and reporting. The availability of computing power and the use of the Internet for many facets of forensic analytics have made all the steps in the process easier. All that is missing now is for forensic investigators, internal auditors, external auditors, and other data analysts to use the methods and techniques on their data.

The first three chapters in the book are an overview of using Microsoft Access, Excel, and PowerPoint for the analysis of data and the reporting of the forensic results. The next nine chapters describe forensic analytic methods and techniques that begin with high-level overviews and then drill deeper and deeper into the data to produce small sets of suspicious transactions. One high-level overview technique reviewed in depth is Benford's Law. Thereafter, two chapters show how correlation and time-series analysis can be used as detective or proactive continuous monitoring techniques. Chapters 15 and 16 discuss, with examples, a forensic risk-scoring technique that would work well in

a continuous monitoring application. Chapter 17 reviews the detection of financial statement fraud. The chapter shows how Benford's Law can be used to detect such frauds and also includes a scoring technique to score divisions for financial reporting fraud. The final chapter reviews the use of forensic analytics to detect purchasing card fraud and possible waste and abuse in a purchasing card environment.

The methods and techniques in the book are discussed and described with results from real-world data. The chapters also include a detailed demonstration of how to run the tests in Access 2007 and Excel 2007. These demonstrations are supported by about 300 screen shots showing the steps used to run the tests. In a few cases, either Access or Excel is demonstrated when that alternative is clearly the way to go. Forensic investigators should have no problem in running these tests in Access 2010 or Excel 2010 using the screenshots in the book.

The companion site for the book is www.nigrini.com/ForensicAnalytics.htm. The website includes the data tables used in the book. Users can then run the tests on the same data and can then check their results against the results shown in the book. The website also includes Excel templates that will make your results exactly match the results in the book. One template is the *NigriniCycle.xlsx* template for all the tests in the Nigrini cycle. The templates were prepared in Excel 2007. The companion site also includes PowerPoint 2007 slides for all 18 chapters. The website also has exercises and problems typical of those found at the end of college textbook chapters. These materials could be used by college professors using the book in a formal college course. With time, more sections will be added to the website and these might include links to useful resources and questions from forensic investigators and my answers to the end-of-chapter questions.

Forensic Analytics is the result of many years of work on forensic analytic projects, starting with my Ph.D. dissertation titled "The Detection of Income Tax Evasion through an Analysis of Digital Distributions." The book was written so that it would be understood by most financial professionals. Ideally, most users will have some experience in obtaining transactional data and some experience with the basic concepts of data analysis, such as working with tables, combining (appending) or selecting (extracting subsets) data, and performing calculations across rows or down columns. Users should understand the basics of either Excel or Access. There are many books covering these basics and also many free resources on the Microsoft website. In addition to the technical skills, the ideal user should have enough creativity and innovation to use the methods as described, or to add twists and tweaks to take into account some distinctive features of their environment. Besides innovation and creativity, the target user will also have a positive attitude and the disposition to, at times, accept that their past few hours of work have all been the equivalent of barking up the wrong tree and after taking a deep breath (and a few minutes to document what was done) to go back (perhaps with new data) and start again. Much of forensic analytics is more like an art than a science and forensic investigators need a personality that matches the iterative process of modifying and refining the tests.

To this day I am still thankful to my Ph.D. dissertation committee for their guidance and supervision of my forensic-based dissertation that was a move into uncharted

waters. I still remember the many Friday afternoon progress sessions with Martin Levy, a professor of Applied Statistics and Quantitative Analysis. A special thanks is also due to the first internal audit directors, Jim Adams, Bob Bagley, and Steve Proesel, that used my forensic analytic services in the mid-1990s. I needed their vote of confidence to keep going. I'd also like to thank the Wiley professionals, Timothy Burgard, Stacey Rivera, and Chris Gage, who turned my manuscript into a quality finished product.

Mark J. Nigrini, Ph.D.
Pennington, New Jersey, USA
February 18, 2011

About the Author

MARK NIGRINI, PH.D., IS an Associate Professor at The College of New Jersey in Ewing, New Jersey, where he teaches auditing and forensic accounting. He has also taught at other institutions, including Southern Methodist University in Dallas, Texas.

Mark is a Chartered Accountant and holds a B.Com. (Hons) from the University of Cape Town and an MBA from the University of Stellenbosch. His Ph.D. in Accounting is from the University of Cincinnati, where he discovered Benford's Law. His dissertation was titled "The Detection of Income Tax Evasion through an Analysis of Digital Distributions." His minor was in statistics and some of the advanced concepts studied in those statistics classes are used in this book.

It took a few years for his work to be noticed by corporate America. The breakthrough came in 1995 when his work was publicized in an article titled "He's got their number: Scholar uses math to foil financial fraud" in the *Wall Street Journal*. This was followed by several other articles on his work and on Benford's Law in the national and international media. A recent article on Benford's Law that discussed Mark's forensic work was published in Canada's *Globe and Mail* on December 22, 2010. Mark has also been interviewed on the radio and television. His radio interviews have included the BBC in London and NPR in the United States. His television interviews have included an appearance on NBC's *Extra*.

Mark has published papers on Benford's Law, auditing, and accounting in academic journals such as *The Journal of the American Taxation Association*, *Auditing: A Journal of Practice and Theory*, *The Journal of Accounting Education*, *The Review of Accounting and Finance*, *Journal of Forensic Accounting*, and *The Journal of Emerging Technologies in Accounting*. He has also published in scientific journals such as *Mathematical Geology* and pure mathematics journals such as the *International Journal of Mathematics and Mathematical Sciences*. Mark has also published articles in practitioner journals such as *Internal Auditor* and the *Journal of Accountancy*. Mark's current research addresses forensic and continuous monitoring techniques and advanced theoretical work on Benford's Law.

Mark has presented many academic and professional seminars for accountants in the United States and Canada with the audiences primarily comprising internal auditors, external auditors, and forensic accountants in the public and private sectors. Mark has presented a number of association conference plenary or keynote sessions with his talk titled "Benford's Law: The facts, the fun, and the future." The release date

of *Forensic Analytics* is planned to coincide with a plenary session to be delivered by Mark at NACVA's Annual Consultants' Conference in San Diego, CA, on June 9, 2011. Mark has also presented seminars overseas with professional presentations in the United Kingdom, The Netherlands, Germany, Luxembourg, Sweden, Thailand, Malaysia, Singapore, and New Zealand. Mark is available for seminars and presentations and he can be contacted at ForensicAnalytics@gmail.com. Other contact information is given on his website www.nigrini.com.

Contents

Preface xi

About the Author xv

Chapter 1: Using Access in Forensic Investigations	1
An Introduction to Access	2
The Architecture of Access	4
A Review of Access Tables	6
Importing Data into Access	8
A Review of Access Queries	10
Converting Excel Data into a Usable Access Format	13
Using the Access Documenter	20
Database Limit of 2 GB	24
Miscellaneous Access Notes	24
Summary	25
Chapter 2: Using Excel in Forensic Investigations	27
Pitfalls in Using Excel	28
Importing Data into Excel	30
Reporting Forensic Analytics Results	32
Protecting Excel Spreadsheets	34
Using Excel Results in Word Files	36
Excel Warnings and Indicators	40
Summary	41
Chapter 3: Using PowerPoint in Forensic Presentations	43
Overview of Forensic Presentations	44
An Overview of PowerPoint	44
Planning the Presentation	45
Color Schemes for Forensic Presentations	46
Problems with Forensic Reports	50
Summary	61

Chapter 4: High-Level Data Overview Tests	63
The Data Profile	64
The Data Histogram	67
The Periodic Graph	69
Preparing the Data Profile Using Access	70
Preparing the Data Profile Using Excel	77
Calculating the Inputs for the Periodic Graph in Access	79
Preparing a Histogram in Access Using an Interval Table	81
Summary	83
 Chapter 5: Benford's Law: The Basics	 85
An Overview of Benford's Law	86
From Theory to Application in 60 Years	89
Which Data Sets Should Conform to Benford's Law?	97
The Effect of Data Set Size	98
The Basic Digit Tests	99
Running the First-Two Digits Test in Access	102
Summary	107
 Chapter 6: Benford's Law: Assessing Conformity	 109
One Digit at a Time: The Z-Statistic	110
The Chi-Square and Kolmogorov-Smirnoff Tests	111
The Mean Absolute Deviation (MAD) Test	114
Tests Based on the Logarithmic Basis of Benford's Law	115
Creating a Perfect Synthetic Benford Set	121
The Mantissa Arc Test	122
Summary	129
 Chapter 7: Benford's Law: The Second-Order and Summation Tests	 130
A Description of the Second-Order Test	131
The Summation Test	144
Summary	151
 Chapter 8: Benford's Law: The Number Duplication and Last-Two Digits Tests	 153
The Number Duplication Test	154
Running the Number Duplication Test in Access	155

Running the Number Duplication Test in Excel	164
The Last-Two Digits Test	167
Summary	172
Chapter 9: Testing the Internal Diagnostics of Current Period and Prior Period Data	173
A Review of Descriptive Statistics	175
An Analysis of Alumni Gifts	178
An Analysis of Fraudulent Data	182
Summary and Discussion	189
Chapter 10: Identifying Fraud Using the Largest Subsets and Largest Growth Tests	191
Findings From the Largest Subsets Test	193
Running the Largest Subsets Test in Access	195
Running the Largest Growth Test in Access	197
Running the Largest Subsets Test in Excel	200
Running the Largest Growth Test in Excel	203
Summary	210
Chapter 11: Identifying Anomalies Using the Relative Size Factor Test	212
Relative Size Factor Test Findings	213
Running the RSF Test	215
Running the Relative Size Factor Test in Access	216
Running the Relative Size Factor Test in Excel	226
Summary	232
Chapter 12: Identifying Fraud Using Abnormal Duplications within Subsets	233
The Same-Same-Same Test	234
The Same-Same-Different Test	235
The Subset Number Duplication Test	236
Running the Same-Same-Same Test in Access	238
Running the Same-Same-Different Test in Access	239
Running the Subset Number Duplication Test in Access	244
Running the Same-Same-Same Test in Excel	248
Running the Same-Same-Different Test in Excel	252

Running the Subset Number Duplication Test in Excel	256
Summary	262
Chapter 13: Identifying Fraud Using Correlation	263
The Concept of Correlation	264
Correlation Calculations	272
Using Correlation to Detect Fraudulent Sales Numbers	272
Using Correlation to Detect Electricity Theft	276
Using Correlation to Detect Irregularities in Election Results	278
Detecting Irregularities in Pollution Statistics	282
Calculating Correlations in Access	287
Calculating the Correlations in Excel	291
Summary	295
Chapter 14: Identifying Fraud Using Time-Series Analysis	297
Time-Series Methods	299
An Application Using Heating Oil Sales	299
An Application Using Stock Market Data	303
An Application Using Construction Data	306
An Analysis of Streamflow Data	313
Running Time-Series Analysis in Excel	319
Calculating the Seasonal Factors	320
Running a Linear Regression	322
Fitting a Curve to the Historical Data	324
Calculating the Forecasts	325
Summary	330
Chapter 15: Fraud Risk Assessments of Forensic Units	332
The Risk Scoring Method	333
The Forensic Analytics Environment	335
A Description of the Risk-Scoring System	336
P1: High Food and Supplies Costs	338
P2: Very High Food and Supplies Costs	339
P3: Declining Sales	340
P4: Increase in Food Costs	342
P5: Irregular Seasonal Pattern for Sales	344
P6: Round Numbers Reported as Sales Numbers	346
P7: Repeating Numbers Reported as Sales Numbers	347

P8: Inspection Rankings	347
P9: High Receivable Balance	348
P10: Use of Automated Reporting Procedures	348
Final Results	349
An Overview of the Reporting System and Future Plans	350
Some Findings	351
Discussion	353
Summary	353
Chapter 16: Examples of Risk Scoring with Access Queries	355
The Audit Selection Method of the IRS	356
Risk Scoring to Detect Banking Fraud	360
Final Risk Scores	364
Risk Scoring to Detect Travel Agent Fraud	364
Final Results	369
Risk Scoring to Detect Vendor Fraud	369
Vendor Risk Scoring Using Access	376
Summary	385
Chapter 17: The Detection of Financial Statement Fraud	388
The Digits of Financial Statement Numbers	388
Detecting Biases in Accounting Numbers	395
An Analysis of Enron's Reported Numbers	398
An Analysis of Biased Reimbursement Numbers	399
Detecting Manipulations in Monthly Subsidiary Reports	404
Predictor Weightings	421
Conclusions	423
Summary	424
Chapter 18: Using Analytics on Purchasing Card Transactions	425
Purchasing Cards	426
The National Association of Purchasing Card Professionals	432
A Forensic Analytics Dashboard	433
An Example of Purchasing Card Data	433
High-Level Data Overview	435
The First-Order Test	438
The Summation Test	440
The Last-Two Digits Test	440

The Second-Order Test	441
The Number Duplication Test	442
The Largest Subsets Test	444
The Same-Same-Same Test	446
The Same-Same-Different Test	446
The Relative Size Factor Test	448
Conclusions with Respect to Card Purchases	449
A Note on Microsoft Office	450
A Note on the Forensic Analytic Tests	451
Conclusion	452

References 455

Index 459

Using Access in Forensic Investigations

FORENSIC ANALYTICS IS THE procurement and analysis of electronic data to reconstruct, detect, or otherwise support a claim of financial fraud. The main steps in forensic analytics are (a) data collection, (b) data preparation, (c) data analysis, and (d) reporting. This book casts a wider net than simply the detection of financial fraud. Using computer-based analytic methods our goal is the detection of fraud, errors, and biases where biases involve people gravitating to specific numbers or number ranges to circumvent actual or perceived internal control thresholds. These analytic methods are directed at determining the likelihood or magnitude of fraud occurring. They would be a part of a fraud deterrence cycle that would include other steps such as employment screening procedures, including background checks. The techniques described in the book rely on the analysis of data, usually transactional data, but at times, other data such as statistical data or aggregated data of some sort.

The main workhorses for the preparation and analysis of data will be Microsoft Access and Microsoft Excel (or Access and Excel, for short). Other valuable and dependable and high-quality tools for data analysis include IDEA, Minitab, and SigmaPlot for preparing high-quality complex graphs. The reporting and presentation of the results is usually done using Microsoft Word and/or Microsoft PowerPoint. These results could include images cropped from various sources (including Access and Excel). Images can be copied and pasted into Word or PowerPoint by using a software tool called Snag-It.

This chapter introduces Access and the components and features of Access that are used in a forensic analytics environment. The next two chapters do the same for Excel and PowerPoint. In summary, Access has almost everything that is needed for a forensic analytics application with reasonably sized data sets, where there is not a high

requirement for high security. Forensic-related applications can be created in Access and other users with little or no knowledge of Access could use the system. The chapter reviews the Access components and features that make it useful for forensic analytics.

AN INTRODUCTION TO ACCESS

Access is Windows-based and so, fortunately, all the basic Windows operations work in Access. Your trusted mouse works just like before with right clicks, left clicks, and double clicks. Access is launched just like any other program using a shortcut or the **Start** button. Copying, moving, naming, and deleting files are done as usual. There are some differences that are mainly related to the fact that Access is a database program that expects the data tables to be continually changed and updated.

Access differs from Word and Excel in that for most users there was no migration from other products. Microsoft did an excellent job in showing people how to do task *x* in Word given that you used to do task *x* following a set of procedures using perhaps WordPerfect or Wordstar. Microsoft also showed people how to do task *y* in Excel given that you used to do task *y* using a series of steps in perhaps Quattro Pro or Lotus 1-2-3. For example, you can still enter `@sum(B1..B5)` in cell **B6** in Excel (2007) and not only will it calculate the sum correctly, but it will convert the formula to = `SUM(B1:B5)` for you. There is no help in Access geared to making you more familiar with the program, because there was not a preceding product that users were used to. This makes the logic of Access a little tricky to follow at first. With practice comes familiarity, and it will not be too long before you will prefer to use Access for those projects that are more suited to Access than to Excel.

One reason for favoring Access over Excel for forensic analytics work is that Access forces some discipline onto the data analysis project. Excel is basically a large free-form rectangle divided into smaller rectangles (called cells). In these cells you can (a) paste images, (b) enter numbers, (c) enter formulas, or (d) display a graph (called a chart in Excel). When you view a number in Excel, unless you click on the cell itself, you are never really sure if this is a data point or the result of a formula (a calculation). Excel is (unfortunately) very forgiving in that a column heading can be repeated (you can call both columns A and B, *People*), Excel does not mind if you call a column *Dollars* and immediately below the field name you enter the word *Rambo*. Excel has some built-in documenting capabilities (including the ability to Insert Comment) but most of the structure and the integrity are left up to the user. Without clear documentation it is easy for another user to have no clue as to what is happening in a complex spreadsheet, and even the original developer might have trouble figuring out what is happening if they look at a complex spreadsheet six months later. The opening screen for Access 2007 is shown in Figure 1.1.

In contrast to Access, most computer programs will at least do something once opened. For example, in PowerPoint you can immediately click on the blank slide and type a title or some text. This is not the case with Access. To get Access to start working

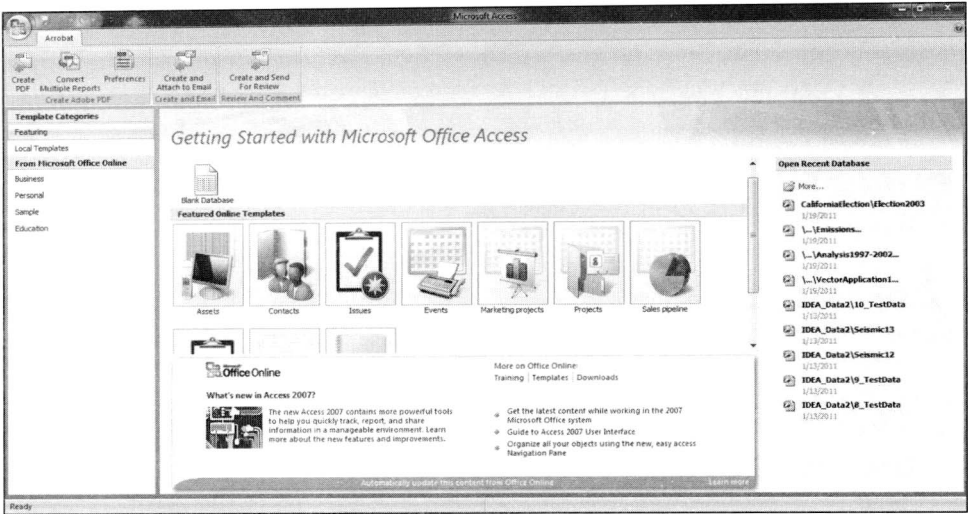


FIGURE 1.1 Opening Screen for Microsoft Access 2007

you either need to open an existing file or you need to create a new blank database. For a new forensic analytics project, the **New Blank Database** is the starting point. Clicking on **Blank Database** will start the series of dialog boxes creating a new Access database. The next step is shown in Figure 1.2.

Figure 1.2 shows the step needed to create an Access database named *Chapter1a.accdb* in a folder named *DataDrivenForensics*. Clicking the **Create** button will give the result in Figure 1.3.

The opening screen of the new database named *Chapter1a* is shown in Figure 1.3. *Table 1* is shown in the open objects panel and this is there so that the spot does not look empty. The table disappears once a new table is created and *Table 1* is closed. The navigation pane on the left lists all the Access objects and the details can be shortened or extended by selecting the drop down arrow and selecting **Object Type** or **All Access Objects**. The architecture of Access and the components of a database are discussed in the next section.

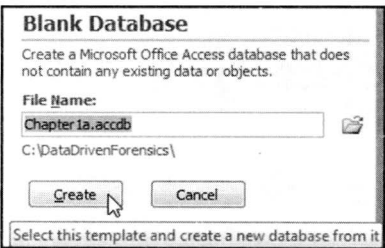


FIGURE 1.2 Creation of a New Blank Database in the *DataDrivenForensics* Folder

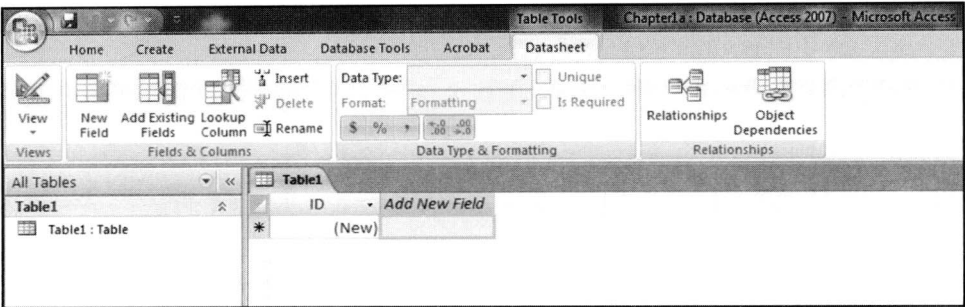


FIGURE 1.3 Opening Screen of a New Access Database Named *Chapter1a*

THE ARCHITECTURE OF ACCESS

The Microsoft Access homepage at <http://office.microsoft.com/en-us/access-help/> has lots of useful and reliable information on Access 2003, 2007, and 2010. The website's opening screen with Access 2007 selected is shown in Figure 1.4.

Extensive Microsoft Access information and help is available as can be seen in Figure 1.4. After selecting the appropriate version on the right (see the arrow in Figure 1.4) the site provides information and help related to using Access. A good starting place, irrespective of your Access version, is the Access Basics section in Access 2010. The *basics* are basically the same for each version except that Access 2007 and Access 2010 use the ribbon for the selection of tasks. There are also other websites with Access information and several of these are listed on the companion site for this book.

An Access database is a tool for collecting, storing, and analyzing data, and reporting information. A database consists of unprocessed data and other objects associated with collecting, editing, adding, deleting, processing, organizing, reporting on, and sharing the data. The objects listed below are of most interest from a forensic analytics perspective:



FIGURE 1.4 Microsoft Website with Access Information and Help