

Using Microsoft Excel for Forensic Analytics

THIS CHAPTER SETS THE stage for our foray into the art of forensic analytics. We will review the three fraud types that are relevant to the analytics tests discussed in this book. We will also look at the main ways in which occupational frauds are discovered. The chapter will also review the main steps in a forensic investigation. The chapter includes a discussion of the main features of Excel 2016 (Excel) that make it a very useful tool for forensic analytics. While Excel 2019 and Excel in Office 365 ProPlus are available at the time of writing, the assumption is that most users are still using Excel 2016, and so the relevant new features are reviewed separately in the supplements to this book.

THE FRAUD TYPES RELEVANT TO FORENSIC ANALYTICS

The FBI defines *fraud* as the intentional perversion of the truth that causes another person or other entity relying on the non-truth to give up something of value. Under their definition fraud schemes are wide-ranging and include credit card fraud, health insurance fraud, identity theft, internet auction fraud, Ponzi schemes, telemarketing fraud, and advance fee schemes. A coverage of all those topics is too much for one book, and so we are going to focus on the three main branches of the fraud tree set out in the highly regarded and often cited *Report to the Nations* published by the Association of Certified Fraud Examiners (ACFE).

The first fraud type is *occupational fraud* (also referred to as *asset misappropriation* or *embezzlement*) in which an employee uses deception to misuse, misapply, or divert an employer's resources or assets to obtain some type of personal gain. This fraud type and the tests that can be run to detect these frauds are the focus of this book. The second fraud type is made up of *bribery* schemes, which involves giving or offering something of value to a person in a position of trust to influence their judgement or conduct. The losses from these types of schemes include reputational losses and the penalties and fines levied by the legal system. The third fraud type is *financial statement fraud* in which a chief financial officer (CFO) or a controller of a company prepares manipulated financial reports (or misrepresented financial data) that is designed to deceive investors, auditors, and analysts about the real financial condition of the organization. The losses from these types of schemes are wide ranging and include not only investors, but also the public at large who lose confidence in the capital markets. Forensic accountants, fraud investigators, internal auditors, and external auditors are all interested in these types of fraud schemes.

The need for analytics tests that are designed to detect embezzlement is evident from two facts. First, the *2018 Report to the Nations* (ACFE, 2018) states that asset misappropriation is the most commonly occurring type of fraud scheme. About 50 percent of these schemes are detected through internal and external tips (40 percent), by accident, notified by law enforcement, and by confession (10 percent). None of these "we were lucky" detection methods are a proactive detective control. The ACFE's data also shows that the fraud schemes that were detected using a proactive detection method, not related to luck, had a shorter fraud duration and a smaller median loss. Second, the FBI's arrest statistics for the past few years have shown a near constant embezzlement arrest rate of 16,000 people per year, interestingly split nearly perfectly 50/50 between males and females. The *2018 Report to the Nations* notes that in only 58 percent of the cases was the perpetrator of an occupational fraud scheme referred to law enforcement. The main reason for not sending the cases to law enforcement was the fear of bad publicity. It seems that we are behind the curve when it comes to proactively going out and detecting fraud using analytical procedures and that occupational fraud is alive and well and thriving in our society. Note also that for-profit companies are not the only victims. Nonprofit organizations such as churches, sports clubs, charities, local governments, and professional associations are all targets for employee embezzlement and for employee and executive conflicts of interest.

The main targets of our analytics will be those types of frauds where the perpetrator invents a series of numbers, usually dollar amounts. The larger the fraud, the larger are those dollar amounts. Some examples of frauds that don't involve number invention are cases where a taxpayer fails to report some of his or her income or a taxable amount (such as a prize in a game show) in a tax evasion case (see Nigrini, 2017). Here an amount of zero dollars is substituted for the correct amount. In the financial statement fraud case described in *USA v. Kumar and Richards* (2004) the executives of a company ("CA") were charged with engaging in a systematic, company-wide practice of falsely and fraudulently recording and reporting, within a fiscal quarter, revenue from software license agreements that were not finalized and signed during that quarter. This practice

was referred to within the company as the “35-day month” or the “three-day window.” This financial statement fraud is one of premature revenue recognition (an issue related to the timing of revenues), which does not involve the invention of any dollar amounts. Finally, in the CIT case, a manager diverted 2,500 end-of-lease computers for his own use, at a loss to the company of some \$637,000 (Michelman, Gorman, and Trompeter 2011). When tangible assets are diverted for the employee’s own use, there is no number invention by the fraudster. In place of forensic analytics aimed at detecting number anomalies, these types of schemes might be detected by an evaluation of internal controls, fraud risk assessments, a diligent scrutiny of budgets and variances, and fraud awareness training.

THE MAIN STEPS IN A FORENSIC ANALYTICS APPLICATION

Data analysis is defined in Wikipedia as the process of inspecting, cleansing, transforming, and modeling data with the goal of discovering useful information, informing conclusions, and supporting decision-making. In internal and external auditing circles, data analysis has moved from the purview of a few IDEA and ACL pioneers to having the Audit Data Analytics Task Force of the American Institute of CPAs releasing an audit guide titled *Guide to Audit Data Analytics* (AICPA, 2017). This was followed by the Chartered Professional Accountants of Canada (CPA Canada) releasing its audit guide titled *CPA Canada Guide to Audit Data Analytics* (CPA Canada, 2018). Drawing on the insights in both publications, the main steps in an audit data analytics (ADA) application are as follows:

1. Set the objectives of the analytics test. These objectives could perhaps be to detect a specific type of employee embezzlement scheme, to detect employees fraudulently accruing points or miles related to the company’s loyalty program, to identify bribery payments, or to identify the risk of a material misstatement in the financial statements related to a management assertion and a class of transactions or account balances.
2. Determine the data type that would be relevant to the test(s). The data type could be financial or nonfinancial data, structured or unstructured data, accounting process and control-related data, historical transactional data, and so on.
3. Select the tests that are best suited for the objective at hand considering data availability. Sometimes the desired data is just not available. For example, a company might want to test whether their checks paid were deposited into a bank account with an exact match between the name of the account holder and the payee on a check. This information is not likely to be available electronically.
4. Select the tools (usually a computer software package) and graphics (again a computer software package) to be used.
5. Obtain the relevant data in the data collection phase (sometimes with difficulty from overprotective human resources or marketing departments); import the data into Excel, Access, IDEA, R, Tableau, or some other software application. Obtaining

the data can be challenging, and from time to time internal auditors and fraud investigators are faced with departmental managers using every stalling trick at their disposal. This stalling phase by the data gatekeepers with the related vague references to privacy laws or claims that the data is “strategically important” is an occupational hazard.

6. As a part of the collection phase the investigator or auditor needs to consider the reliability of the data used. Issues here include whether the data is from an internal or an external source, whether the related controls are effective, and whether the data was obtained directly from the source. In one application the author had no choice but to accept disbursements data from the senior purchasing manager in an Excel format where one of the objectives of the analytics tests was to ascertain whether the disbursements manager himself was committing fraud. It would have been an easy matter for the manager to delete his own fraudulent transactions from the Excel table before sending it over to the team for analysis.
7. The data preparation phase includes data cleansing or data scrubbing. This involves the detection and correction (or removal) of corrupt or inaccurate records from the data tables. This step identifies incomplete, incorrect, inaccurate, or irrelevant parts of the data and the replacement, modification, or deletion of some of the data. An example of data cleansing with pollution statistics data involved (a) removing subtotals, as is commonly needed with government and financial statement data, (b) removing countries with zero or very low emissions since they were not really relevant to the research question, (c) making some geographic adjustments to keep the data consistent despite changes in national borders (e.g., by adding East and West Germany for 1988 and 1989 and naming the country “Germany”), and (d) reconfiguring the data in a relational database table format with the familiar constructs of fields and records (columns and rows).
8. Running the data analysis tests or performing the ADA follows the data cleansing phase. Here the tests described in Chapters 2 through 16 are applied. These tests generally use summaries, calculations, groupings, comparisons, and visualizations. An important point here is that if the initial results of the ADA test indicates that aspects of its design or performance needs to be revised, then the changes should be made and the tests rerun. It is easy to imagine a research scientist at a pharmaceuticals company running and rerunning tests before the *Eureka moment*. The same generally applies to running forensic analytics tests.
9. The results are now evaluated, and the user concludes whether the objectives set at the outset have been accomplished. If this is not the case, the forensic accountant should plan and perform different procedures or use different data sources or data periods to accomplish the objectives.
10. The results of the analysis are reported to a select audience. The results could include tables, graphs, charts, other visualizations, and summaries of selected records. In a forensic setting care needs to be taken to allow only the level of visibility that is appropriate for the audience, and to prevent the audience from changing the contents of the reports.

Not all the steps apply to every forensic analytics application. For example, purchasing card data is almost always reliable and will need very little, if any, data cleansing or data scrubbing. The results of the various tests will make varying uses of summaries, calculations, groupings, comparisons, and visualizations. It is expected and hoped that with the passage of time the use of an ADA would become more effective and efficient. The learning curve would involve learning from prior experiences. It should be expected that the analytics process is an iterative one with a continual decline in the number of false positives. A false positive is a transaction, a payment, or a financial statement line item that has been identified as a fraud or an anomaly when the item is legitimate and authorized. These false positives use valuable time and effort on the part of the forensic accountant. PwC (2018) report that many of the users of technology to combat fraud and/or economic crimes believed that their tests were producing too many false positives.

THE FINAL REPORT

Forensic analytics work done in Excel, Access, IDEA, Tableau, R, or any of the other data analysis options will inevitably be included in a forensic report. Examples of this type of work might be a net worth analysis, the evaluation of an insurance claim, listings of suspicious transaction, listings of fraudulent transactions, and perhaps graphs showing average prices paid for some item or perhaps income comparisons with other employees. A forensic report would include the findings, conclusions, recommendations, and corrective actions taken. The various headings might be as follows:

1. Executive Summary (the scope of the investigation, the objective, the approach taken, and a summary of the key findings)
2. Scope
3. Approach
4. Findings
5. Summary
6. Financial Impact of the Fraud
7. Recommendations
8. End of Report
9. Exhibits

The forensic analytics work could be included in the Findings section and/or the Exhibits. Fortunately, Excel results can be sent to Word documents reasonably easily. The options for copying results to Word are discussed later in this chapter. The Summerford Accountancy's fraud vulnerability assessment prepared for the Wake County Board of Education is available for download from the companion website. Another example, the internal Toshiba investigation into accounting irregularities, is also available for

download from the companion website. A useful set of sample fraud examination documents is available on the website of the ACFE at <https://www.acfe.com/acfe-twocolumn.aspx?pageid=6764&terms=>.

AN OVERVIEW OF EXCEL

Excel is a software program that allows us to populate a rectangular grid, called a worksheet, with numbers, text characters, and images. Readers interested in some Excel history, its features, and some technical details can read the Microsoft Excel page on Wikipedia (https://en.wikipedia.org/wiki/Microsoft_Excel). For our purposes it is enough to say that Excel can perform many tasks and most users probably use less than 10 percent of the program's functionality. Some of the main tasks that can be done with Excel in a forensic analytics environment are:

- Importing and accessing data from databases and government and corporate websites.
- Storing data in an easily retrievable format.
- Performing calculations related to forensic analytic tests.
- Grouping data and calculating statistics (such as sums or counts) on a per-group basis.
- Creating graphs that give visual insights into transactional data.
- Copying the results easily and seamlessly into Word and PowerPoint.

The fact that Excel has about the same numbers of users worldwide as Instagram (if we don't split hairs about whether we're talking about daily or monthly active users) supports the statement that Excel is an excellent product. However, Excel is still limited to data sets with a maximum of 1,048,576 rows and 16,384 columns. The row and column limits are 2^{20} and 2^{14} for anyone that might be interested in the source of the numbers. The row limit presents a problem in some cases. Data tables of earth science data can easily exceed 1 million records and, for accounting environments where the transactional amounts are small (e.g., coupon redemptions for a large consumer goods company), the number of records can easily exceed 10 million. For environments that are likely to exceed this limit, it is therefore best to use Access, IDEA, or R for the analysis. Also, using Excel in a 32-bit environment makes it subject to a 2-gigabyte (GB) memory limit. So essentially Excel only works well with many rows and a few columns, or with many columns and only a few rows. How many such rows and columns will depend, in part, on whether the data includes long text fields and how the numbers are formatted.

Excel users *en masse* love the flexibility of the program. The cells in a spreadsheet can contain labels, data, calculated fields, graphs, text characters, images, and almost anything else that can make it onto the clipboard. The result can include named ranges, complex IF formulas, and filters giving us a system of dependent calculations that link together to form the final product. This has the drawback that it is difficult for someone else (or even the spreadsheet creator) to know what is happening in the spreadsheet

when they open it a few months later. The term *spreadsheet risk* is used to describe the risk of faulty decisions being made based on errors that have crept into spreadsheets. Readers are encouraged to read some of the horror stories in the *Horror Stories* page of the European Spreadsheet Risks Interest Group (<http://www.eusprig.org/index.htm>). Highly interested readers could even consider attending one of their conferences. Analytics programs such as Access and SQL Server improve transparency by having tables, queries, and reports that are clearly labeled and separated. Those programs also have comprehensive documentation capabilities for the database objects.

Linked to the transparency issue is the fact that Excel mixes the source data and the results of the analysis. Access keeps the tables, queries, and reports separate, and as such we always know whether we are looking at the source data or the results of calculations. As processes become more complex and more users want more information from the same Excel tool, it is inevitable that users will try to upgrade the spreadsheets to do just one more task. This will involve more creativity and even more complexity, and the additional new results will become interrelated, adding to spreadsheet risk. Access allows users to add more tables, queries, and reports in a systematic way without disturbing the current calculations and relationships. Access databases can more easily accommodate additions and changes precisely because of the separation of tables, queries, and reports, and the required format of the database and its tables.

IMPORTING DATA INTO EXCEL

The normal situation in a forensic analytics environment is that transaction or balance data is imported into Excel. This is usually an easy matter by using either the **Data→Get External Data** series of steps or the **Copy** and **Paste** sequence. In any analytics project we need our data set to be complete to get the desired results. It is a rather serious flaw to be working with incomplete data. For this reason, we should be very suspicious of Excel data tables with exactly 16,384 or 65,536 rows. An example is shown in Figure 1.1.

A (blurred) purchasing card dataset with 16,384 records is shown in Figure 1.1. Data sets with exactly 16,384 or 65,536 records are created when the accounting system housing the data source exports the data in an Excel format and that system is set to export only 16,384 or 65,536 records because that is the Excel row limitation that is programmed into the system. A limitation also comes into play when a **Copy** and **Paste** sequence of commands (that is, using the Clipboard) is used to copy data from Access to Excel.

The first solution to the row limit problem is to have the source system download the data as a text file. The fields can be comma delimited (a .csv file) if none of the field values include commas (e.g., Bachmai International, Inc.). If the field values do include commas, then tab delimited fields or fixed width fields should work fine. The second solution to the row limit problem is to import the data into Excel using the **Data→Get External Data** commands followed by the relevant options.

There is a little issue with the Connections feature in Excel. The **Data→Get External Data** commands imports (say) Access data into Excel. This process will import all the records that can fit into an Excel file (up to 1,048,576 rows). The import

	A	B	C	D	E
16364	STEP	Q03	10.65	10/29/2018	
16365	STEP		11.07	10/31/2018	
16366	STEP		305.30	09/09/2018	
16367	STEP	0	25.71	09/25/2018	
16368	STEP	OR	168.17	09/29/2018	
16369	STEP		243.04	09/30/2018	
16370	STEP	AB#	13.85	10/02/2018	
16371	STEP	SN	62.56	10/02/2018	
16372	STEP	AB#	51.58	10/03/2018	
16373	STEP	2	8.87	10/06/2018	
16374	STEP	AS	289.48	10/07/2018	
16375	STEP		22.80	10/10/2018	
16376	STEP	AB#	28.32	10/14/2018	
16377	STEP		22.80	10/16/2018	
16378	STEP		165.90	10/16/2018	
16379	STEP	SN	121.32	11/03/2018	
16380	STEP		78.98	11/07/2018	
16381	STEP	E	32.23	11/07/2018	
16382	STEP	JTH	1500.00	11/10/2018	
16383	STEP	JTH	1201.47	11/17/2018	
16384	STEP		8.27	11/20/2018	
16385					
16386					
16387					
16388					
16389					
16390					

FIGURE 1.1 A Table with Exactly 16,384 Records

step also formats the data as an Excel table. Two further steps are needed in a forensic analytics environment. First, the data needs to be “disconnected” from the source so that the Excel file is not updated or changed. This disconnection step would use **Data→Connections→Connections** followed by clicking the **Remove** button. Excel will give a warning, and the step is finished with **OK** (to acknowledge the warning) and **Close**. While the table format is pleasing to the eye, it does create some issues for calculated fields because they now “point” to the original source data. It is easier to work with the data when it is not in this table format. The steps to undo the table format starts by selecting any cell in the table (e.g., A2) and then using a right click and the **Table→Convert to Range** sequence of commands.

SOME USEFUL EXCEL FORMATTING FEATURES

The first useful Excel formatting option for complex spreadsheets is the cell styles option, where cells can be color coded according to their contents. Selected cell style options have been applied to a Benford’s Law first digits (see Chapter 3) table in Figure 1.2.

Figure 1.2 shows the *Bad* cell style applied to the Mean Absolute Deviation value in F12, the *Warning Text* style applied to the label in E12, the *Output* style applied to the counts in column B, and the *Calculation* style applied to C3:G11. Using these styles

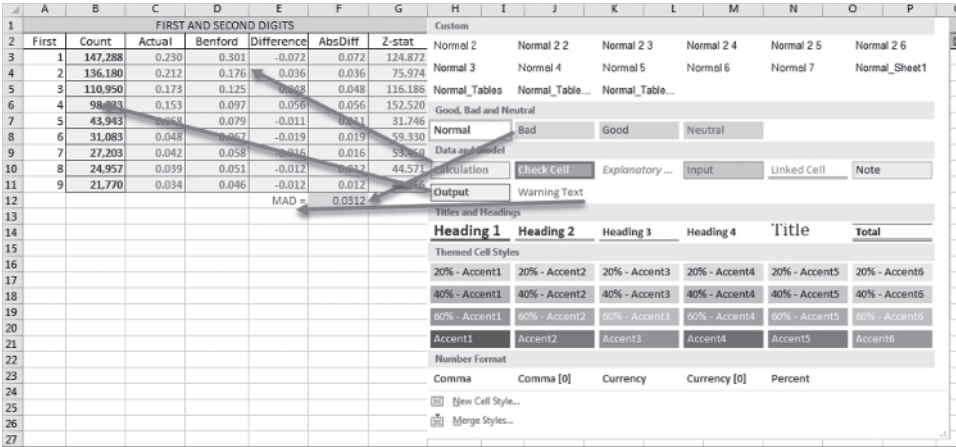


FIGURE 1.2 Excel Cell Styles Applied to a Benford’s Law Table.

makes it easier for the creator to see what is happening and for others to understand what is happening in the spreadsheet, if they know what the styles mean. The **Data and Model** and the **Titles and Headings** styles allow users to prepare a worksheet with a very professional and polished look to it with built-in visual guidance for users.

The second useful Excel formatting option for complex spreadsheets is the conditional formatting option, where cells can be color coded according to various criteria. Conditional formatting is applied using **Home**→**Styles**→**Conditional Formatting**→**Top/Bottom Rules**→**Top 10 Items**, which returns the dialog box shown in Figure 1.3.

Figure 1.3 shows some conditional formatting applied to the top values in the range (G2:G10). The options for conditional formatting are nearly limitless and include being able to use a formula to determine which values to apply a special format to (see the **More Rules** option). This feature is particularly useful to identify the three largest Z-Statistics, which relates to the digits with the largest deviations from the expected Benford proportions.

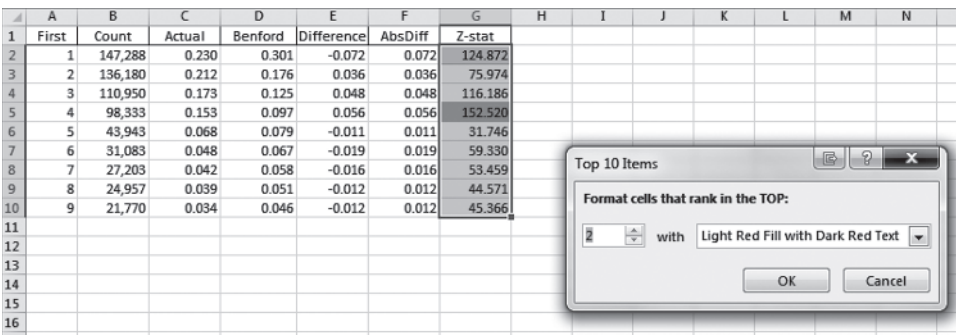


FIGURE 1.3 Conditional Formatting Applied to the Z-Statistics in a Benford’s Law Table.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	1	147,288	0.290	0.301	-0.072	0.072	124.872		Light							
2	2	136,180	0.212	0.176	0.036	0.036	75.974									
3	3	110,950	0.173	0.125	0.048	0.048	116.186									
4	4	98,333	0.153	0.097	0.056	0.056	152.520									
5	5	43,943	0.068	0.079	-0.011	0.011	31.746									
6	6	31,083	0.048	0.067	-0.019	0.019	59.330									
7	7	27,203	0.042	0.058	-0.016	0.016	53.459									
8	8	24,957	0.039	0.051	-0.012	0.012	44.571									
9	9	21,770	0.034	0.046	-0.012	0.012	45.366									
10									Medium							
11																
12																
13																
14																
15																
16																
17																
18																
19																
20																
21																
22																
23									Dark							
24																
25																
26																
27																
28																
29																
30																
31																

FIGURE 1.4 Table Formatting Applied to a Benford’s Law Table.

The third useful Excel formatting option for complex spreadsheets is to convert the range to a **Table**. This is done using **Home**→**Styles**→**Format as Table** to give the result shown in Figure 1.4.

Figure 1.4 shows the Table Style *Medium 2* applied to a Benford’s Law table. Not only is this pleasing to the eye, but the results can be filtered using the drop-down menus in each of the header cells. The usual forensic analytics applications will use one of the **Number Filters** options. For example, applying the number filter “Greater Than” and then entering zero will show a table with only three rows related to the digits 2, 3, and 4. This is logical since the data (discussed at length in Chapter 4) has been seeded with fraudulent journal entries in the \$2,000 to \$4,999 range. One important filtering caveat is that if you enter a formula `=SUM(A:A)` in cell H2, the answer will be 45. Excel will sum all the values in column A and not just the ones that are visible to you after applying a filter.

PROTECTING EXCEL SPREADSHEETS

Confidentiality of case-related data is especially important in a forensic analytics environment. Excel offers several protection options but none of these safeguards rivals Fort Knox, and forensic accountants should understand that physical control over the Excel file is just as important as the file-based protections.

The first protection option is to protect the Excel file itself. This is done with a password by clicking on **File** and then selecting **Save As**. The next step is to select **Tools**→**General Options** at the bottom section of the dialog box. Excel will then prompt you to create a password to open the file and a password to modify (edit) the file. The step to get to the password creation dialog box is shown in Figure 1.5.

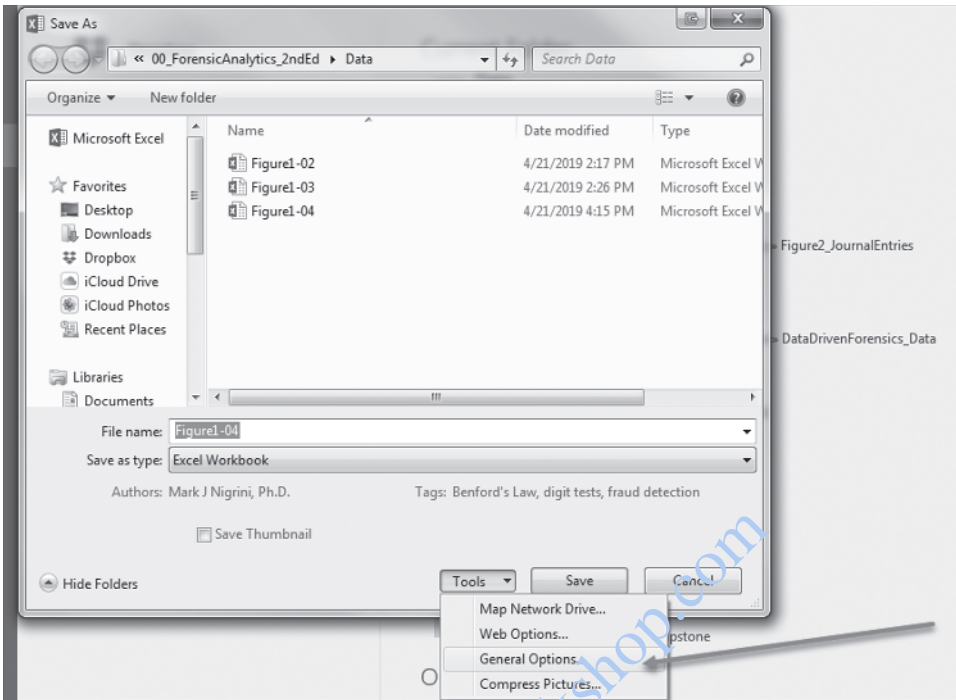


FIGURE 1.5 The Path to the Passwords Dialog Box for the Excel File.

Figure 1.5 shows the step used to access the dialog box used to create a password to open and to modify the file. Excel warns users that the passwords cannot be “recovered” before saving the file. If the file is saved, and later reopened, Excel indicates that the file is protected and requires the password to open the file. Password protected files can still be accessed by a determined and tech savvy person, and the password is only a good protection against most (but not all) people. Excel also has other levels of protection. These are activated through the **Review**→**Changes** group and includes protecting the sheet, protecting the workbook, and allowing users to edit ranges.

THE VALUABLE “IF” FUNCTION

Some of Excel’s functions are particularly valuable for identifying forensic-related data anomalies. The valuable IF function and the general form of this function is as follows:

$$=IF(\text{Logical Test}, \text{Value if TRUE}, \text{Value if FALSE})$$

where the *Logical Test* is one that can be evaluated as being either true or false. The logical test can itself be composed of several within-equation equations using perhaps the AND or OR options. The general usage of the true/false options will be to code the cell (that is the “answer” to the question) with a value of 1 if the condition is true and either a 0 or a blank (shown by two successive quotes " and " as "").

	A	B	C	D	E
1	Date	Name	Amount	Merchant	Description
2	25-Apr	Andrew Davis	\$1,531.40	SHI CORP	SOFTWARE/LICENSE
3	27-Apr	Andrew Davis	\$559.28	DMI* DELL K-12/GOVT	8 MEMORY KITS
4	10-Apr	Andrew Davis	\$417.35	FLORIDA ARF/ RESPECT	COLOR CARTRIDGES
5	20-Apr	Andrew Davis	\$14.60	1800GOFEDEx 10010007	SHIPPING
6	4-Apr	Maria Martinez	\$4,000.00	OFFICE FURNITURE&DESIGN	PC CUBICLE MOVE
7	28-Apr	Maria Martinez	\$345.20	SHRED IT NEWARK	1T SHRED SERVICE
8	18-Apr	Maria Martinez	\$93.60	D J*WALL ST JOURNAL	SUBSCRIPTION
9	20-Apr	Maria Martinez	\$23.90	AMAZON MKTPLACE PMTS	CABLES
10	5-Apr	Maria Martinez	\$2.99	AMAZON MKTPLACE PMTS	DATA CABLE
11	14-Apr	Amy Powell	\$1,095.00	DOCS - DEANNE'S OFFICE	LABOR FOR BOARDS
12	28-Apr	Amy Powell	\$660.25	SAP AMERICA-TRAINI	PC#100130
13	24-Apr	Amy Powell	\$54.98	BEST BUY 00004358	CASE FOR PHONE
14					

FIGURE 1.6 Purchasing Card Data Sorted by Name Ascending and by Amount Descending.

The Relative Size Factor (RSF) test (described in Chapter 7) is the ratio of the largest value in a subset divided by the second largest value. This test uses two clever IF functions to calculate these ratios. Let's assume that we only have three employees with purchasing cards and each person has from three to five transactions in the month. After sorting the data by employee name ascending and by dollar amount descending (our data preparation step) we have the data set that is shown in Figure 1.6.

The first step in running the RSF test would be to make a copy of the sheet using a right click on the worksheet name tab at the bottom of the worksheet and clicking **Move or Copy** and then creating a copy. The next step is to delete the columns that are not needed (A, D, and E). A counter is then created in Column C so that we can identify the largest and the second largest *Amount* for each employee. This is done by entering a "1" (without the quotes) in C2 and then by starting the counter in C3. A screenshot of the IF function creating the counter is shown in Figure 1.7.

	A	B	C	D	E	F
1	Name	Amount	Counter			
2	Andrew Davis	\$1,531.40	1			
3	Andrew Davis	\$559.28	2			
4	Andrew Davis	\$417.35				
5	Andrew Davis	\$14.60				
6	Maria Martinez	\$4,000.00				
7	Maria Martinez	\$345.20				
8	Maria Martinez	\$93.60				
9	Maria Martinez	\$23.90				
10	Maria Martinez	\$2.99				
11	Amy Powell	\$1,095.00				
12	Amy Powell	\$660.25				
13	Amy Powell	\$54.98				
14						

FIGURE 1.7 The First Step in Calculating the Relative Size Factor.

	A	B	C	D	E	F	G	H
1	Name	Amount	Counter	RSF				
2	Andrew Davis	\$1,531.40	1	=IF(AND(C2=1,C3=2),B2/B3,"")				
3	Andrew Davis	\$559.28	2					
4	Andrew Davis	\$417.35	3					
5	Andrew Davis	\$14.60	4					
6	Maria Martinez	\$4,000.00	1					
7	Maria Martinez	\$345.20	2					
8	Maria Martinez	\$93.60	3					
9	Maria Martinez	\$23.90	4					
10	Maria Martinez	\$2.99	5					
11	Amy Powell	\$1,095.00	1					
12	Amy Powell	\$660.25	2					
13	Amy Powell	\$54.98	3					
14								

FIGURE 1.8 The Second Step in Calculating the Relative Size Factor.

The counter is started by entering a "1" in C2 because this is the first record of the first employee. The logic of the counter function in Row 3 is that if the employee name is unchanged ($A3=A2$), then the counter must be incremented by 1 ($C2+1$); and if the employee name has changed, then start at 1 again (1 is the third term in the function). The next step is to calculate the RSF only for those rows where the counter equals 1 and where the counter on the next row is a 2. We do not calculate an RSF for employees (subsets) with only one record because they essentially don't have a second largest number. This calculation is done using an IF function in Column D, and this is shown in Figure 1.8.

We need two conditions to be true for us to calculate a valid RSF. We must be on a row where the counter value is 1 and where the next counter value is 2. We do this with the AND function nested in the IF function with the conditions $C2=1$ and $C3=2$. If an employee only has one transaction, that row will have a counter value of 1; but the next row will also have a counter value of 1, and so the RSF will not be calculated. Since the conditions are true for Row 2, the RSF is calculated using $B2/B3$. Where the conditions are not true, the formula will give us a blank cell because of the double quotes in the third term of the IF function (""). When the formula is copied down to D13, the result will be one RSF for each of the three employees and nine blank rows. The final step is to highlight cells C2:D13 and to use **Copy** and **Paste Special** → **Paste Values** to convert the formulas to values. This is important because a listing of just the RSFs will change the order of the rows, and these formula results are highly dependent on the order of the rows. We want to be able to sort the data using the RSF values that have been calculated. Formatting the RSF values to be numeric with 2 decimal places will give us tidier-looking results, and sorting the RSF from largest to smallest will let us focus on the large values at the top of the output table. The result is shown in Figure 1.9.

	A	B	C	D	E
1	Name	Amount	Counter	RSF	
2	Maria Martinez	\$4,000.00	1	11.59	
6	Andrew Davis	\$1,531.40	1	2.74	
11	Amy Powell	\$1,095.00	1	1.66	
14					

FIGURE 1.9 The Final Table of Relative Size Factors.

Figure 1.9 shows the results, and in this application high RSF values could signal errors in that the largest amount was incorrect (maybe it should have been \$400 and not \$4,000), or the largest amount was incorrectly attributed to that specific subset. We used the AND function because we needed the row value to be 1 and the next row's value to be 2. In other cases we might need only one of (say) two conditions to be true, in which case we would use Excel's OR function. For any Excel function a Google search will show you examples of how to use these functions. The suggestion is that you look at the well-written Microsoft Office Support pages first, and if that doesn't work, then try one of the other sites with Excel help.

THE PIVOTTABLE ROUTINE

PivotTables are very useful for forensic analytics tests. PivotTables allow us to summarize and analyze data in a way that allows us to see patterns and trends. One useful test is the largest growth test (see Chapter 6), where we look for subsets that have experienced large growth on a period-on-period basis. These subsets could be employee travel claims, purchases from a vendor, employee overtime hours, customer allowances granted by an employee, or various types of refunds. The period in question could be a year, a half-year, or a quarter. In line with the previous example let's assume that we only have three employees with purchasing cards and each person has from four to six transactions over a two-year period. After sorting the data by employee name and by date, we have a 2018–2019 worksheet that is shown in Figure 1.10.

To calculate the growth in spending on a per person basis from 2018 to 2019 using the data in Figure 1.10, we need to create a field that shows only the year. This is done using Excel's YEAR function, and the formula and the results are shown in Figure 1.11.

Using the YEAR function (`=YEAR(A2)`) in cell D2 in Figure 1.11 gives us the year of the transaction, namely 2018. When this formula is copied down to the last row of the data, we have column D correctly indicating the year as either 2018 or 2019. To calculate the growth in spending we need the total for each employee for each of the two years. The command sequence to get the PivotTable dialog box visible starts by first putting your cursor anywhere in the data "rectangle" that will be analyzed. In the example in Figure 1.11 the cursor is at D2 which is in the A1:D17 range. The commands that follow are **Insert**→**Tables**→**PivotTable**→**OK**. The next step is to use the PivotTable Fields section on the right of the newly created sheet in the spreadsheet to

	A	B	C	D	E	F
1	Date	Name	Amount			
2	2-Oct-2018	Amy Powell	\$1,095.00			
3	29-Dec-2018	Amy Powell	\$660.25			
4	5-Jun-2019	Amy Powell	\$54.98			
5	11-Dec-2019	Amy Powell	2029.06			
6	26-Dec-2019	Amy Powell	1295.00			
7	19-Mar-2018	Andrew Davis	\$1,531.40			
8	15-Oct-2018	Andrew Davis	\$559.28			
9	13-Aug-2019	Andrew Davis	\$417.35			
10	25-Oct-2019	Andrew Davis	\$14.60			
11	9-Dec-2019	Andrew Davis	1791.78			
12	7-Feb-2018	Maria Martinez	\$345.20			
13	22-Oct-2018	Maria Martinez	\$4,000.00			
14	15-Dec-2018	Maria Martinez	\$93.60			
15	14-Feb-2019	Maria Martinez	275.25			
16	5-Jun-2019	Maria Martinez	2718.12			
17	13-Nov-2019	Maria Martinez	5569.40			

FIGURE 1.10 Purchasing Card Data Sorted by Name Ascending and by Date Ascending.

D2 : X ✓ fx =YEAR(A2)						
	A	B	C	D	E	F
1	Date	Name	Amount	Year		
2	2-Oct-2018	Amy Powell	\$1,095.00	2018		
3	29-Dec-2018	Amy Powell	\$660.25	2018		
4	5-Jun-2019	Amy Powell	\$54.98	2019		
5	11-Dec-2019	Amy Powell	2029.06	2019		
6	26-Dec-2019	Amy Powell	1295.00	2019		
7	19-Mar-2018	Andrew Davis	\$1,531.40	2018		
8	15-Oct-2018	Andrew Davis	\$559.28	2018		
9	13-Aug-2019	Andrew Davis	\$417.35	2019		
10	25-Oct-2019	Andrew Davis	\$14.60	2019		
11	9-Dec-2019	Andrew Davis	1791.78	2019		
12	7-Feb-2018	Maria Martinez	\$345.20	2018		
13	22-Oct-2018	Maria Martinez	\$4,000.00	2018		
14	15-Dec-2018	Maria Martinez	\$93.60	2018		
15	14-Feb-2019	Maria Martinez	275.25	2019		
16	5-Jun-2019	Maria Martinez	2718.12	2019		
17	13-Nov-2019	Maria Martinez	5569.40	2019		

FIGURE 1.11 Purchasing Card Data with a Year Indicator in Column D.

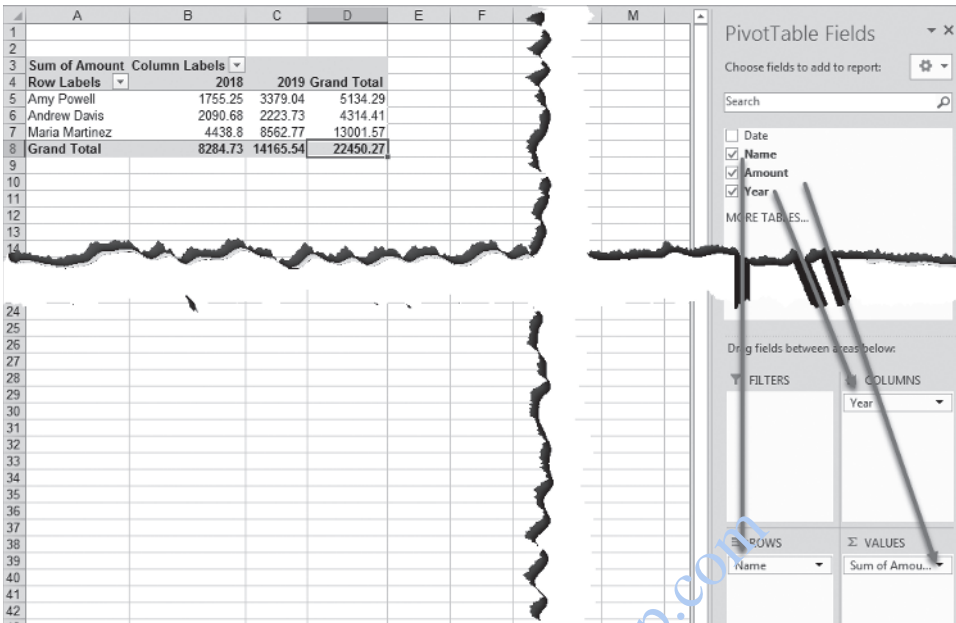


FIGURE 1.12 PivotTable Results Showing the 2018 and 2019 Totals for the Three Employees.

grab (with a left click) and to drag the fields to the right places in the *Filters*, *Columns*, *Rows*, and *Values* sections at the bottom right. The result after dragging *Year* to *Columns*, *Name* to *Rows*, and *Amount* to *Values* is shown in Figure 1.12.

The PivotTable results in Figure 1.12 show the total purchases for both 2018 and 2019 for each of the three employees. This information can now be used to calculate the percentage change from 2018 to 2019. The results (calculations not shown) are increases of 92.5, 6.4, and 92.9 percent respectively. An analysis of these increases would focus on the large percentage increases while paying less attention to those increases that were from a small base (e.g., \$100 increasing to \$400).

THE VALUABLE VLOOKUP FUNCTION

The VLOOKUP function is used when we add a column to our data in which Excel looks up a numeric value or a text description. As a simple example we might want to classify counties as *Small* if the population is under 25,000 people and *Large* if the population is 25,000 or more people. Excel would then look at the population number (e.g., 120,585) and then “consult” the small table with the classifications and return *Large* as the result. Using the same logic, an accounting professor could use VLOOKUP to assign grades (A+, A, A-, B+, B-, and so on) to students based on their average percentages.

In Excel we can extract the day of the month, the month, and the year from a date field using the DAY, MONTH, and YEAR functions. The WEEKDAY function can

	A	B	C	D	E	F	G	H
1	Date	Name	Amount	Quarter		Month	Quarter	
2	21-Jan-2020	Amy Powell	\$1,095.00			1	1	
3	9-Mar-2020	Amy Powell	\$660.25			2	1	
4	7-Apr-2020	Amy Powell	1295.00			3	1	
5	3-May-2020	Amy Powell	2029.06			4	2	
6	13-May-2020	Amy Powell	\$54.98			5	2	
7	4-Jan-2020	Andrew Davis	\$1,531.40			6	2	
8	4-Mar-2020	Andrew Davis	\$559.28			7	3	
9	17-Apr-2020	Andrew Davis	\$14.60			8	3	
10	31-May-2020	Andrew Davis	\$417.35			9	3	
11	7-Jun-2020	Andrew Davis	1791.78			10	4	
12	10-Jan-2020	Maria Martinez	\$93.60			11	4	
13	20-Jan-2020	Maria Martinez	\$4,000.00			12	4	
14	28-Jan-2020	Maria Martinez	\$345.20					
15	3-May-2020	Maria Martinez	275.25					
16	30-May-2020	Maria Martinez	2718.12					
17	1-Jun-2020	Maria Martinez	5569.40					

FIGURE 1.13 Purchasing Card Data for Q1 and Q2 of 2020.

even identify the day of the week (Sunday, Monday, and so on). However, Excel (unlike Tableau) cannot identify the quarter (1, 2, 3, or 4) from a date field. However, the quarter can be identified using the VLOOKUP function. The hypothetical purchasing card data is used to demonstrate the VLOOKUP function and the data had been changed to pertain to two quarters (as opposed to two years) and is shown in Figure 1.13.

The goal of this analytics exercise is to classify each purchase as belonging to either Quarter 1 (Q1) or Quarter 2 (Q2). We would then calculate the percentage increase by using a PivotTable as was done in the previous section. We therefore want Excel to identify the month of the purchase and then to refer to the small table in F1:G13 and then to assign the purchase to Q1 or to Q2 (there are no purchases for Q3 or Q4). The VLOOKUP command for D2 is as follows:

```
=VLOOKUP (MONTH (A2) , $F$1 : $G$13 , 2 , FALSE)
```

The formula instructs Excel to calculate the *Month* using column A and then to go to the small table on the right (in F2:G12) and then to return the quarter number from the second column in the small table. The FALSE part of the formula instructs Excel to find an exact match. The small table on the right is identified using \$F\$1:\$G\$13 so that the formula can simply be copied from D2 to D3:D17 and it will still be pointing to the correct range. The results are shown in Figure 1.14.

Figure 1.14 shows that the VLOOKUP function has correctly classified each purchase as being in either Q1 or Q2. The totals for the two quarters can be calculated using a PivotTable as is shown in the preceding section. The same result could be obtained by using a complex set of nested IF functions. These would start with
`=IF (MONTH (A2) <=3 , 1 , IF (MONTH (A2) <=6 , 2 , IF (MONTH (A2) <=9 , 3 , 4)))`. Long formulas such as that can easily contain errors, and it is always difficult at the end to figure out how many closing brackets are needed. Besides errors, this formula

=VLOOKUP(MONTH(A2),\$F\$1:\$G\$13,2,FALSE)								
	A	B	C	D	E	F	G	H
1	Date	Name	Amount	Quarter		Month	Quarter	
2	21-Jan-2020	Amy Powell	\$1,095.00	1		1	1	
3	9-Mar-2020	Amy Powell	\$660.25	1		2	1	
4	7-Apr-2020	Amy Powell	1295.00	2		3	1	
5	3-May-2020	Amy Powell	2029.06	2		4	2	
6	13-May-2020	Amy Powell	\$54.98	2		5	2	
7	4-Jan-2020	Andrew Davis	\$1,531.40	1		6	2	
8	4-Mar-2020	Andrew Davis	\$559.28	1		7	3	
9	17-Apr-2020	Andrew Davis	\$14.60	2		8	3	
10	31-May-2020	Andrew Davis	\$417.35	2		9	3	
11	7-Jun-2020	Andrew Davis	1791.78	2		10	4	
12	10-Jan-2020	Maria Martinez	\$93.60	1		11	4	
13	20-Jan-2020	Maria Martinez	\$4,000.00	1		12	4	
14	28-Jan-2020	Maria Martinez	\$345.20	1				
15	3-May-2020	Maria Martinez	275.25	2				
16	30-May-2020	Maria Martinez	2718.12	2				
17	1-Jun-2020	Maria Martinez	5569.40	2				

FIGURE 1.14 Purchasing Card Data for Q1 and Q2 of 2020 with a Quarter Indicator in Column D.

would be difficult for someone else to audit or update. They would, for example, need to understand that Excel works through the IF statements sequentially and acts as instructed the first time that a condition is found to be true.

Note that if a date is dragged to the Columns section of a PivotTable, then Excel can quite easily parse the data into quarters (*Qtr1*, *Qtr2*, *Qtr3*, and *Qtr4*). The quarters lookup was simply used to show how the VLOOKUP function works. The VLOOKUP function might be useful for companies that don't close their books on the last day of a quarter (January 31, June 30, September 30, and December 31) but might instead use 13-week quarters with an occasional 14-week quarter to catch up to our 365-day year.

USING EXCEL RESULTS IN WORD FILES

Tables and graphs can be copied from Excel and used as exhibits in forensic reports. They could also be copied and used to create tables and figures in generic research reports. For this section we will assume that the data in A1:D17 in Figure 1.14 needs to be copied into a Word document. The steps would start by highlighting the Excel data range that needs to be copied to the Word document. The next Excel step would be to click **Home**→**Clipboard**→**Copy**. Then return to the Word document and place the cursor where the table should be pasted. The paste command in Word is **Home**→**Clipboard**→**Paste** and **Paste Special** to get to the dialog box shown in Figure 1.15.

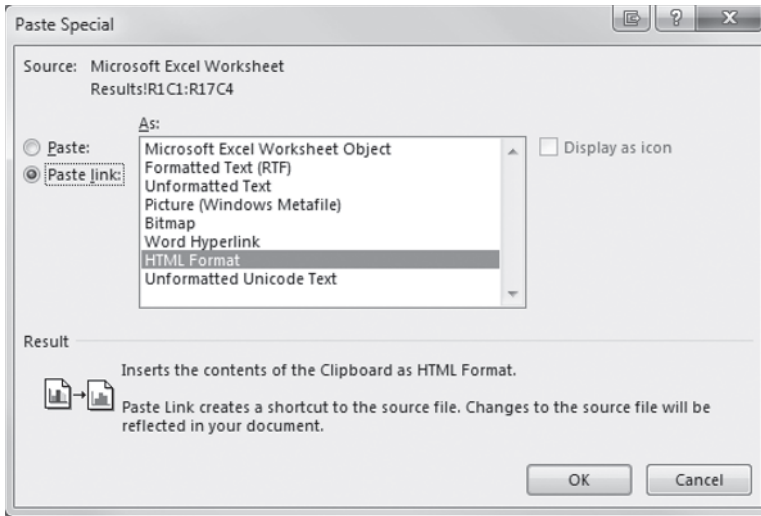


FIGURE 1.15 The Options for Pasting an Excel Table into a Word Document.

The dialog box used to paste an Excel table into a Word document is shown in Figure 1.15. If the **Paste link** and the **Microsoft Office Excel Worksheet Object** options are chosen, all changes in the original Excel file are copied to the Excel table in the Word document. This could be useful if the dollar amounts were to change. A change to the Excel file would be reflected in the Word document. To paste the table into Word without any updating we would use:

- **Paste** (the radio button above **Paste link**) and
- **Microsoft Office Excel Worksheet Object**

The exhibits in a forensic report could include an Excel graph (or chart) embedded in the Word document. The procedure would start by selecting the chart in Excel. The **Copy** command in Excel is activated using **Home**→**Clipboard**→**Copy**. After switching to Word, the **Home**→**Clipboard**→**Paste**→**Paste Special** sequence would give the various options. The options that should be chosen are:

- **Paste** and
- **Microsoft Excel Chart Object**

The chart can now be updated from within Word by double-clicking the embedded chart in the Word document, and then clicking anywhere within the chart. A hatch-marked border will appear around the chart, and the embedded object will appear in an Excel workbook window. The object can now be edited by using the Excel chart-editing tools in Word. To end the editing simply click outside the chart anywhere in the Word document.

It is suggested that you not link the tables and graphs and keep control over when and how the Word document gets updated. It is easy to imagine that you, or someone else, opens the Excel file and makes some changes, perhaps experimenting with some What-If scenarios without realizing that a Word document (a forensic report nonetheless) is being unknowingly updated in the background. The suggestion for your reports is to use Snagit (<https://www.techsmith.com/>) to copy images to Word documents. Snagit needs to be purchased, and if your budget is tight, then the Snipping Tool that comes with Windows will also do a good job (see <https://support.microsoft.com/en-us/help/13776/windows-10-use-snipping-tool-to-capture-screenshots>).

EXCEL WARNINGS AND INDICATORS

While working in Excel, users are almost sure to spot a little green triangle in the top left corner of a cell. In the example shown in Figure 1.16 the error is that there is an inconsistent formula. In each case the USD amount is the transaction amount (Column C) divided by the conversion rate for USD in H2. For transactions already in USD they must simply be divided by 1. Cells E4 and E13 have the conversion factors with no reference to the transaction amounts in column C. The data, the error indicators, and the error message is shown in Figure 1.16.

The amounts in cells E4 and E13 in Figure 1.16 are errors because they display the exchange rate and not the transaction amount divided by the exchange rate. For help on the error indicator you need to click on the small down arrow between the exclamation point and 18.94. The fact that **Inconsistent Formula** is highlighted tells you that the error is an inconsistent formula. This is a real error and the formula needs to be corrected. If the formulas were indeed correct, then the step to remove this minor eyesore (the green triangle) is to select the down arrow to the right of the

	A	B	C	D	E	F	G	H
1	Date	Name	TransAmount	Currency	USD		Currency	Rate
2	21-Jan-2020	Amy Powell	1,095.00	USD	1095.00		USD	1.00
3	9-Mar-2020	Amy Powell	660.25	USD	660.25		CAD	1.35
4	7-Apr-2020	Amy Powell	1,295.00	MXP	18.94		MXP	18.94
5	3-May-2020	Amy Powell	2,029.06	USD				
6	13-May-2020	Amy Powell	54.98	USD				
7	4-Jan-2020	Andrew Davis	1,531.40	USD				
8	4-Mar-2020	Andrew Davis	559.28	USD				
9	17-Apr-2020	Andrew Davis	14.60	USD				
10	31-May-2020	Andrew Davis	417.35	USD				
11	7-Jun-2020	Andrew Davis	1,791.78	USD				
12	10-Jan-2020	Maria Martinez	93.60	USD	93.60			
13	20-Jan-2020	Maria Martinez	4,000.00	CAD	1.35			
14	28-Jan-2020	Maria Martinez	345.20	USD	345.20			
15	3-May-2020	Maria Martinez	275.25	USD	275.25			
16	30-May-2020	Maria Martinez	2,718.12	USD	2718.12			
17	1-Jun-2020	Maria Martinez	5,569.40	USD	5569.40			

FIGURE 1.16 The Green Triangles in the Top Left Corners Are Error Indicators.

exclamation point and to click **Ignore Error**, which is the fourth option in the explanation box. The **Ignore Error** selection should not be made before considering whether there might be a real problem with a formula requiring perhaps an analysis of the precedents and dependents. You can control the messages that you receive by using **File→Options→Formulas** and selecting or deselecting the error-checking rules of your choice.

EXCEL DASHBOARDS

A recent *Wall Street Journal* article noted that “When Walt Bettinger’s 3 a.m. alarm sounds, among the first things the Charles Schwab chief executive does is check how much new money his company has pulled in over the past 24 hours.” If that was me, the first thing that I’d check would be the clock to see whether it really was 3 a.m.! However, it seems a safe bet that the new money numbers and several other key metrics would be presented on a management dashboard. Analytics-based dashboards provide ever-changing and updated graphical and tabular views of the key performance indicators related to an objective or a business process. That is, the graphs and tables on display are like the dashboard of your car that tells you, amongst other things, how fast you are going and how long you might be able to keep this up without stopping to fill up with gasoline. For an easy-to-understand overview of dashboards you can search for “dashboards for Power BI designers,” and then you can scan the various docs.microsoft.com pages that are returned by the search.

The types of dashboards of interest to us are more closely related to continuous monitoring applications. Here the organization wants *alerts* related to suspicious transactions (possible errors or fraud) and continuous controls monitoring where we want to see whether our internal controls are operating as intended. Other relevant monitoring applications might be related to anti-money laundering activities and sanctions screening (activities related to avoiding doing business with sanctioned entities).

Excel’s dashboard functions, impressive as they are, are more geared towards data scanning in that it slices (with a slicer) and dices our data every which way, and it is up to us to spot an adverse trend or ratio and relate this to some type of fraudulent activity or anomaly. This fraud-detection approach is quite dependent on the time period that we have under review and the dimensions (a Tableau term) that we have available for slicing and dicing. If we only have 2019 and 2020 data, we can’t possibly know trend-wise how the amounts compared to 2017 and 2018. Dimensions are essentially data categories and the commonly used dimensions are related to people, products, places, and times. For example, we might be looking at the customer refund percentages per region for our four regions (northeast, northwest, southeast, and southwest) and we might completely miss the sky-high refund percentage in City A because it’s lumped together with 100 other cities in the region. Similarly, we might want to review insurance claims approved just below authorization levels, and our claims paid data might simply not have the adjuster’s name or an adjuster identifier.

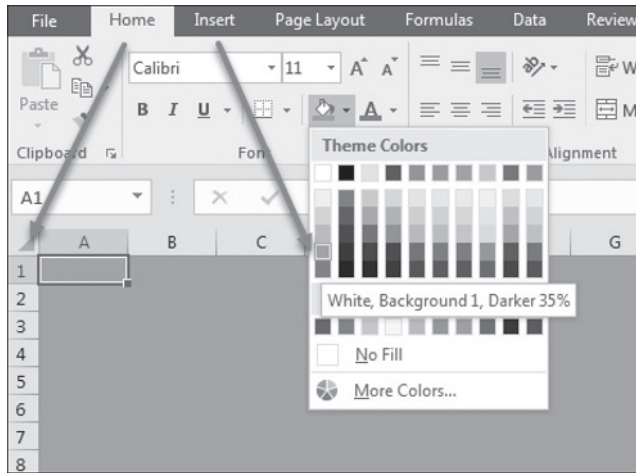


FIGURE 1.17 Creating the *Dashboard Sheet*.

We will create a simple dashboard in this section, and interested readers can refer to the small mountain of information on dashboards available on Google and YouTube. Interestingly, an Excel dashboard is not some special sheet but is rather a normal sheet that has had a little color added to make the individual cells and their margins no longer visible. We will use the source data shown in Figure 1.11 and the PivotTable shown in Figure 1.12. Our first step is to create a new sheet for our dashboard by clicking the + sign to the right of the sheet names at the bottom of our worksheet to give us a new sheet (*Sheet1*), which we will rename with a right click to *Dashboard*. The next step is to select the whole sheet by clicking the right-angled triangle icon to the left of the column indicator A and above the row indicator 1 (essentially the top-left corner of the spreadsheet). This is shown by the leftmost with an arrow in Figure 1.17. The next step is to fill the worksheet, and my preference is White, Background 1, Darker 35%. The steps are shown in Figure 1.17.

We are now in a position to create a graph (chart) that is linked to our PivotTable (a PivotChart). This is done by placing the cursor inside the (tidied-up) PivotTable (say B5) and then using **Insert**→**Charts**→**PivotChart**→**PivotChart**→**OK** to accept the clustered column. The result after some resizing is shown in Figure 1.18.

We will now add some horsepower to our Excel tool, namely the developer tools and the camera. We can add the Developer tab using **File**→**Options**→**Customize Ribbon**. Now under **Customize Ribbon**→**Main Tabs**, select the **Developer** check box. Click **OK**. The Code and the Add-ins tabs are the ones that are most useful for our purposes.

To add the camera in its own customized tab requires a visit to **File**→**Options**→**Customize Ribbon** and then selecting **Camera** under **Customize Ribbon**→**All Commands**. At this stage it gets a bit tricky to add a *New Group*. An easier way to get access to the camera is to simply add the camera to the Quick Access Toolbar on the top left of the screen. Click the drop down box at the top left of the screen which should show **Customize Quick Access Toolbar** and now select **More Commands**.

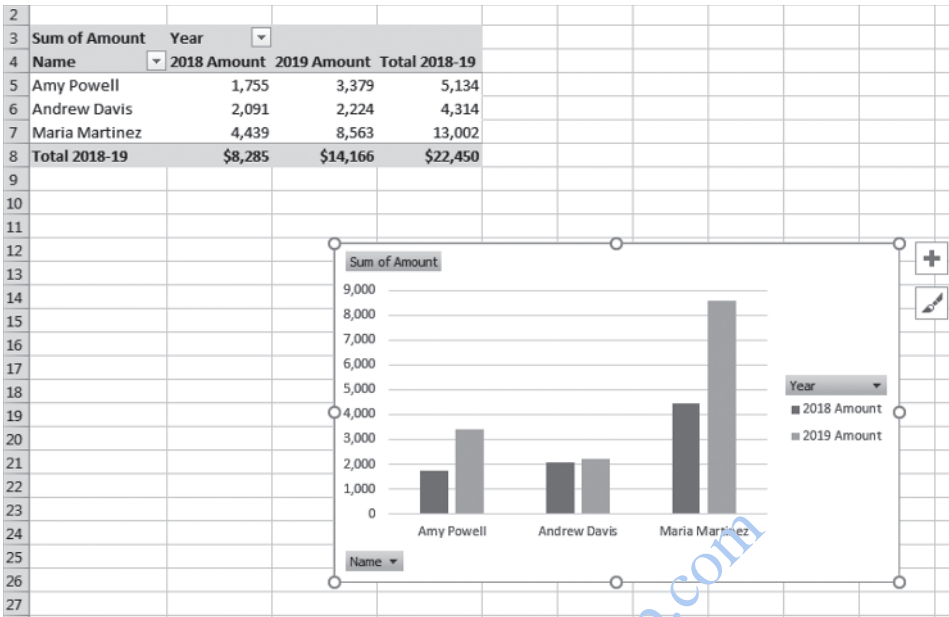


FIGURE 1.18 The PivotChart Based on the PivotTable

Use the **All Commands** option and **Add the Camera**. Click **OK**. You should now see a camera icon on the quick access toolbar.

We will now create a basic dashboard with our purchasing card PivotTable totals, the PivotChart graph, and two slicers. The PivotTable was made more professional looking by tidying up the headings and the column widths and formatting. To view the various totals in the dashboard we need to highlight the PivotTable and then click on the Camera icon, move over to the Dashboard worksheet, and click **Paste**. You should now see the PivotTable on the *Dashboard* sheet.

The next step with our simple dashboard is to move the PivotChart to the **Dashboard** worksheet. Highlight the PivotChart. Use **Design**→**Location**→**Move Chart** and then the radio button to get it as an **Object in** the Dashboard worksheet. After a little tidying up and resizing, our dashboard should look like the dashboard shown in Figure 1.19.

Our next step is to place two slicers on our dashboard. These slicers will allow us to slice and dice our data. Make sure that your cursor is in any cell on the original PivotTable (not the camera image on the dashboard). Slicers are added using **Insert**→**Filters**→**Slicer**. Check **Name** and **Year** and click **OK**. Now move the slicers to the dashboard using **Cut** and **Paste**. The result should be as is shown in Figure 1.20.

By clicking on the names in the *Name* slicer, the graph will show the data for any single employee, any group of employees, or all the employees combined (use the Shift key to highlight all the names). Similarly, by clicking on the years in the *Year* slicer, the graph will show the data for 2018, or 2019, or both years combined. As the graphs change, so too will the source data in the PivotTable.

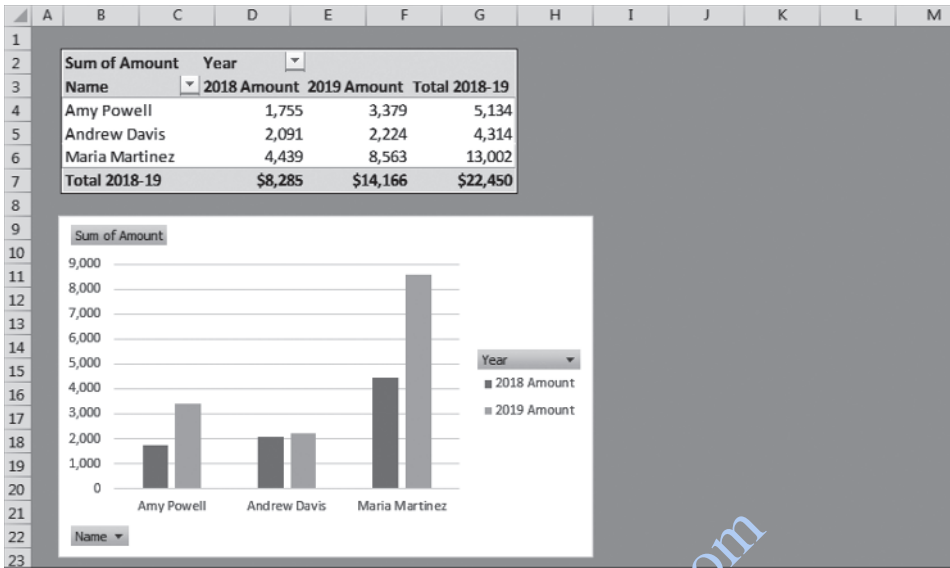


FIGURE 1.19 The Purchasing Card Dashboard with Employee Names and the Total Spend Amounts.

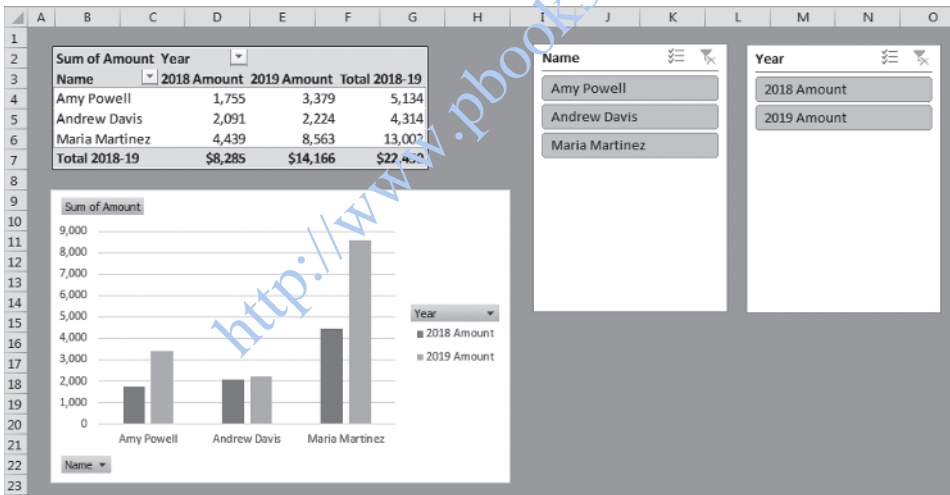


FIGURE 1.20 The Purchasing Card Dashboard with Employee Names, Total Spend, and Two Slicers.

DASHBOARDS IN PRACTICE

In real-world accounting systems transactions are recorded using journal entries. Whereas students in Accounting 101 might make one journal entry for the total sales for a month, or for the wages for a month, the enterprise resource planning (ERP) systems used by companies use journal entries for each individual sales transaction and journal entries for each shift worked by its employees. This results in potentially millions

of journal entries for the year. The auditing standards of the Public Company Accounting Oversight Board (PCAOB) note that intentional financial statement misstatements are carried out by initiating inappropriate journal entries or by adjusting amounts that should be reported in the financial statements on spreadsheets. The standards require the auditor to design procedures to test the appropriateness of journal entries and spreadsheet adjustments. There are two types of journal entries, the first type being *automated* (or automatic) journal entries, which automatically flow from sales, manufacturing production, and purchases transactions. The second type are manual entries that are used for adjustments, accruals and prepayments, funds transfers, internal billings, cost allocations, reversals of prior entries (voids), closing entries, and special events such as stock dividends or acquisitions. The manual entries are under the control of top management and are the type that could be used for financial statement fraud.

To assist in their audits in general, and in their scrutiny of journal entries, PwC (formerly known as PricewaterhouseCoopers) use a tool called *Halo for Journals* which relies heavily on a dashboard-based user interface. The main dashboard and some of the drill-down options can be seen at <http://halo.pwc.com/>. The main dashboard gives an overview of the number of journal entries, some descriptive statistics, the financial statement line items with the largest number of underlying entries, an automated/manual breakdown, and statistics for the Top and Bottom 5 users. The testing tool gives their auditors a host of options such as duplicates and reversals, large entries, unusual days, and unusual times. The third dashboard allows users to apply various filters isolating manual or automatic journal entries that have a specific financial statement line item impact (e.g., manual entries greater than \$100,000 impacting the cash and cash equivalents line item). These tests all presumably make use of filters and slicers such as those shown in Figure 1.20.

An interesting interactive dashboard is available on the website of the General Services Administration (GSA). This organization's SmartPay Program provides government agencies and other organizations with accounting services related to general purchases, travel expenses, and fleet expenses. An overview of the organization can be found by clicking on the various links on their homepage at <https://smartpay.gsa.gov/#a1>. An example of an interesting dashboard can be found at <https://spdatawarehouse.fas.gsa.gov/spdw/>. Users can click on either the *Program Data* or the *Business Lines* tabs on the left, and several options are available on the right side of the dashboard. Users can select Spend (total amounts), Transaction counts, Card counts, and average transaction amounts. In the bottom right quadrant users can drill down to view the monthly trend for a specific agency. Readers should click on the *Text View of Program Statistics* link near the top left of the page. This link will download the source data in an Excel format. A review of the source data will allow users to understand what is needed to "fuel" a business intelligence dashboard.

SUMMARY

Forensic analytics is the procurement and analysis of electronic data to reconstruct or detect fraud, errors, or biases. The main steps in a forensic analytics project are data collection, data preparation, data analysis, and reporting. The three fraud types that are most relevant to forensic analytics are embezzlement schemes, bribery schemes,

and financial statement fraud schemes. About one half of frauds are discovered by internal and external tips, by accident, by a law enforcement notification, and by confession. This shows that there is a real need for proactive fraud detection activities. Such activities would start off by setting the objectives of the tests. Thereafter the relevant data would be accessed using an effective and efficient set of software tools. The actual analytics tests would be run after the data has been prepared for the analysis. The final phases would involve evaluating the results and reporting the results to a select audience.

Excel is the software tool of choice for most forensic accountants. Excel works very well when it comes to importing or connecting to data. Excel can store the data in an easily retrievable format. The program can easily be used for calculations and grouping of records and for creating insightful graphs. There are some caveats related to Excel and these relate mainly to (a) a limit on the number of records, (b) a lack of transparency between data, formulas, and the results of formulas, (c) the difficulty in adding yet more functions and reports to an existing application, and (d) the ability to use Excel in multiuser environment. Even with these issues, Excel is still the favorite among accounting and auditing professionals.

Importing data into Excel is usually not a problem. Excel can import data seamlessly from Access and a host of other database programs. For data imported from Access it is best to remove the connection to Access (so that the Excel file does not update when the Access tables are changed) and to convert the Excel table to a normal range.

Excel has many data formatting options that help in interpreting and visualizing our data. Data bars resembling a very small graph can shade the cells, allowing users to interpret the relative sizes without reading the numbers. Excel offers other cell styles such as *good*, *bad*, and *neutral* and other formatting options that might be useful in a forensic accounting context.

Excel has some data protection options that include the option to password protect the Excel file. A second level of protection is a second password to write or change any cell contents in the workbook. Other levels of protection include the ability to hide rows or columns, or even the entire worksheet. Cells can also be protected with read only privileges. The password protections are useful against an average Excel user, but they can be circumvented by a determined and tech savvy person.

The results of tests and techniques performed in Excel are often included in forensic reports prepared using Word. These results could be extracts from Excel files and graphs and charts. Excel content can be easily copied over to Word documents and several copying options are available. The Excel data can be linked to the Word document, in which case the Word document is updated (changed) when the source data is changed. The Excel data can also be copied and pasted, in which case the Word document stays unchanged even when the source file is changed. Graphs and charts can also be copied from Excel to Word.

Excel's IF function and the PivotTable routines form the foundation for the types of calculations that are relevant to forensic analytics. The VLOOKUP routine allows us to search for and display information in a table. In a table of state populations VLOOKUP will allow us to look up the population of (say) West Virginia. In another context this function will also let us assign a letter grade (A, B, and so on) to a student based on their average percentage in a class.

Despite the reservations related to having data, formulas, the results of formulas, and text and images all potentially on the same worksheet, Excel remains a valuable tool for forensic analytics. The use of a dashboard could help users to see some key metrics at a glance. These metrics are updated when the underlying data is updated.

<http://www.pbookshop.com>