



**Spreadsheets:
A High-Risk Tool for Data Analysis**

W H I T E P A P E R



■ ■ ■ INTRODUCTION

When the electronic spreadsheet was first introduced to the business world in the late 1970s and early 1980s, it became an almost immediate success. The usefulness, adaptability and power of spreadsheets proved to be a strong draw and, arguably, helped launch the broad-based adoption of personal computers in business. Today, electronic spreadsheets are ubiquitous – and for good reason. They have become an indispensable tool for those involved in financial calculations and modeling.

A recent article on AccountingWEB.com, quoting a survey of financial executives, revealed that 92 percent of all public companies use spreadsheets for critical accounting activities. Uses varied from revenue accounting entries to revenue scheduling, allocation and redistribution. The article contends that spreadsheets are widely used because many key revenue recognition and reporting tasks are still not fully automated in Financial and ERP systems.

Spreadsheet capabilities – calculations, conditional statements, linking and programming through macros – make them extremely adept at creating ad hoc applications. They can be applied to any number of areas including budgeting, inventory, financial modeling, and data entry of financial information and reporting.

Paradoxically, what makes spreadsheets so attractive is at the root of their shortcomings. Organizations need to balance the ease-of-use, flexibility and low procurement cost of spreadsheets against their shortcomings for use within an enterprise-level context. These shortcomings may be categorized in three main areas:

- Lack of data integrity – values may be altered deliberately or accidentally
- Error prone – errors in input, logic, data interfaces and use
- Not in line with standard IT regimes for critical applications – documentation, testing and version control

It is clear that the shortcomings of spreadsheets may lead to questioning the quality of the data they produce. Poor data quality results in inaccurate reporting and misinformed decision-making. Thus, the reliance on spreadsheets in financially material applications results in an unacceptable level of business and regulatory risk.

The reality of the situation is that many organizations using spreadsheets for key applications are doing so out of necessity. It isn't always practical or feasible to develop, manage or implement custom applications or to procure point solutions to fill the gap that spreadsheets address. Since a spreadsheet application is better than reverting to manual processes, spreadsheets will continue to be used. In order to deal with the risks introduced by spreadsheets, management needs to take greater care in how and where spreadsheet applications are used and audit needs a means of independently evaluating the accuracy of the data and reports that the organization depends on.

■ ■ ■ RISKS INHERENT IN SPREADSHEET USE

Spreadsheet errors are as pervasive as spreadsheets themselves. A *Computer World* article of May 2004 indicates that up to 91 percent of recently audited spreadsheets contained errors. In *The Journal of Property Management* of July 2002, it was reported that between 30 and 90 percent of all spreadsheets suffer from at least one significant user error.

The impact is staggering and the stories are legendary. The following are but a few examples that come from government and industry across the globe:

"...data entry error of \$118,387" ¹

"...hefty \$11 million severance error was traced to a faulty spreadsheet" ²

"...firm makes \$30M spreadsheet error." ³

"...misstated by \$644M: undetected spreadsheet error in 'ad hoc' process" ⁴

The range and magnitude of errors depends on a number of factors. There may be simple errors in spreadsheet formulas or in the coding of macros. Sometimes they are attributed to simple user error – something as basic as the copy and paste function. The more complex the spreadsheet becomes, the higher the probability of a material error. Errors may be introduced through inappropriately defined cell ranges, improperly referenced cells or cross-linking. Many of these errors are completely invisible to the end user. Macros imbedded in spreadsheets may never have been documented, tested or placed under version control. Data imported or copied into the spreadsheet may be inadvertently changed, altered or converted. These risks are primarily accidental in nature, though a countless number of cases exist where fraud was being perpetrated through the clever manipulation of spreadsheets.

Despite the errors, spreadsheets are used consistently within many domains and audit is no exception. Using a tabular approach to data analysis in an auditing context is persuasive and given the flexibility and accessibility of spreadsheets, it is not surprising that we find spreadsheets in such wide use. The audit function, however, has unique sensibilities and requirements.

Requirements of the Audit Function

Auditing is an independent, objective assurance and consulting activity designed to add value and improve an organization's operations. It helps an organization accomplish its objectives by bringing a systematic, disciplined approach to evaluate and improve the effectiveness of risk management, control, and governance processes. ⁵

In a financial context, audit typically delivers an independent assessment of an organization's accounting and reporting practices, business risks, and internal controls. They establish audit coverage by determining how well internal controls mitigate the risks of material misstatement in the financial statements. To validate their judgments, the audit function may drill down to the level of an organization's transactions and analyze that data using substantive testing methods.

¹ <http://www.columbiatribune.com/2006/Feb/20060222News009.asp>

² <http://www.marketwatch.com/News/Story/Story.aspx?guid=%7B276FD56F%2D00CA%2D42AE%2DAD70%2DC66DF571FC77%7D&dist=rss&siteid=mktw>

³ <http://www.abc.net.au/news/newsitems/200506/s1394937.htm>

⁴ <http://www.gao.gov/atext/d04754t.txt>

⁵ Institute of Internal Auditors Professional Practices Framework http://www.theiia.org/index.cfm?doc_id=123

In a statutory context, auditors provide organizations with assurance on the truth and fairness of an organization's financial information. They give expert advice on controls and processing system weaknesses and provide confirmation of accounting treatments with respect to complex transactions. In today's increasingly complex regulatory environment, auditors also address regulatory reporting requirements such as those under Sarbanes-Oxley Section 404.

Many auditors have sensibly adopted technology solutions in place of manual processing to aid in these tasks. While no single software tool can manage all of the audit tasks, data analysis is one area in which technology solutions can be very effective. Data analysis software can help increase the breadth, depth and frequency of audits. It can help increase the productivity of the audit department and enhance the value they provide to their organizations.

Like any customer, auditors can choose either a general purpose and easily available tool such as spreadsheets, or a purpose-built tool. The manifest advantage of a purpose-built tool is that it addresses the very real needs of the users - in this case, auditors.

AUDIT AND DATA ANALYSIS

When evaluating a data analysis technology for audit, there are a number of essential attributes that should be considered. These may be divided into five areas:

1. Data accessibility
2. Audit-specific capabilities
3. Logging and automation
4. Suitability for enterprise-class deployment
5. The company behind the product

Data Accessibility

Simply accessing the data can be a daunting task for auditors. This is due to the data request process that most organizations have in place. A typical scenario is as follows:

- a. An individual determines a need for a report or a data extract for analytical purposes
- b. A request for the data is made to the IT department specifying type of data, information required, time period to cover and output format
- c. IT puts the request into their schedule
- d. Data is received two to three weeks later
- e. Initial analysis identifies that either
 - i.) Some information is missing – (account number, postal code, vendor ID) or
 - ii.) Additional information is needed that was not anticipated when the first data request was made
- f. An additional request is sent to IT for supplemental data and the waiting period begins again

The above scenario is not uncommon and is at odds with both the internal and external audit process. Firstly, auditors do not have the luxury to wait several weeks for data extracts. Audit cycles cannot accommodate weeks of waiting for data. Secondly, there is usually some programming or report writing conducted by IT in order to accomplish the data extraction. During this process it is likely that some sort of aggregation or filtering is done. As a result, critical information may be inadvertently omitted. Therefore, the integrity of the data must be called into question. As a result, both ad-hoc and ongoing analysis objectives can be compromised.

Under pressure to do more in less time, auditors are looking to eliminate obstacles and streamline audit processes. An effective audit technology enables auditors by providing them with direct data access. This has the joint benefit of streamlining the overall audit process and relieving busy IT staff from repeated data requests by the audit function. Simply allowing audit to “self serve” the data they need for their audits is insufficient. Volume, variety and veracity are also highly material when considering the suitability of an analysis technology for the audit process.

Volume

An effective data analysis technology for audit must be able to analyze entire data populations to ensure that the entire picture is visible. While statistical sampling has an important place in audit processes, greater assurance is realized by analyzing 100 percent of data populations to ensure anomalies, errors and exceptions are identified. Analysis of entire data populations allows for unprecedented insight into business operations. Suspicious transactions may be detected sooner and corrective action initiated before problems escalate, become material weaknesses or require external reporting.

According to recent research by Gartner Group, massive volumes of data are being accumulated by organizations at an ever-increasing rate – growing by as much as 30 percent annually. In many cases, there is simply too much data to consider downloading or importing the data to a PC platform for analysis. An effective data analysis solution in today’s environment needs to incorporate client/server solutions that provide a robust and dependable technical architecture. In such a solution, data can be analyzed in place, eliminating the need for data extraction, reducing network traffic and minimizing the risks involved in converting and duplicating critical business data.

By comparison, large data files easily overwhelm typical spreadsheet applications. They do not feature client/server architectures and require data to be duplicated, transferred and imported onto PC platforms for analysis.

Variety

Most organizations rely on a variety of applications that run on a variety of operating systems, collecting data in a variety of formats. An effective data analysis tool for the audit process needs to be able to read and compare a broad variety of data formats including relational data, legacy data, spreadsheets, report files, flat files, XML and XBRL-formatted data. Where data resides in databases, an optimal solution will have the ability to directly access and analyze data without having to extract and duplicate the data set. A comprehensive solution will have server components that can run on a variety of operating systems and directly interface with a variety of databases in order to cover the different server platforms deployed in your organization.

While spreadsheets have recently become more adept at importing data, they still fall short of being able to deal with data from different formats and operating environments. The risk is the inadvertent modification of the data during the conversion process. For instance, mainframe data is usually in EBCDIC format and cannot be read by a PC-based spreadsheet without conversion.

Veracity

Veracity, or the truthfulness or accuracy of data is paramount in the audit process. A data analysis technology for audit must ensure data integrity and quality. With data extracts and format conversions, the integrity of data can be inadvertently compromised. A suitable audit technology will perform read-only operations on source data – either from flat files or direct database interfaces. Where possible, such applications leave the data in place and analyze it at the source, thereby avoiding the risks associated in converting or duplicating critical business information.

Sometimes a seemingly benign task, such as sorting data, can lead to error. Many spreadsheet users have been tripped-up by sorting just some of the columns in a spreadsheet, resulting in garbled and meaningless data. If the spreadsheet is subsequently saved, recovery is impossible. An audit-oriented technology needs to protect the user from accidentally changing values and the integrity of the records in the data set. It must preserve the veracity of the data to prevent the skewing of analysis results, which could lead to material errors in findings and erroneous audit recommendations.

Spreadsheets can deliberately or accidentally change the data values and representations once imported into their format, without record.

Audit-Specific Capabilities

A data analysis technology for audit needs to have the features and functionality that auditors require to do their job effectively. Not only should it deal with the data access challenges, it also needs to support the way in which auditors work and the types of analytics that are appropriate to the audit task.

Some aspects of data analysis in audit involve assessing the integrity of business processes and practices, evaluating the efficacy of controls, conducting risk assessments and in some cases fraud detection. Invariably this means that data must be analyzed from a diversity of sources to seek patterns and relationships. Users need to organize their view of the company data in a way that suits the analytical objectives. This view gives users the ability to set an appropriate context from which to compare and contrast data from diverse sources. An effective solution for this type of work allows users to visually represent these diverse sources of information from a unique point-of-view and organize the view of the data to match their working style.

For instance, if part of an audit process is looking for indicators of fraud, one might compare an employee master file with an approved vendor database. If there is a match between an employee's address and the address of a vendor, it might indicate the presence of a "phantom vendor" and that an employee is attempting to perpetrate fraud. In such a case, the user is well suited to have a technology that allows them to visually present these data files in relationship to one another. If there are other elements required for this task, such as references to audit procedures, specific index files, or libraries of filters or formulae and test scripts, the user should be able to organize these too in an appropriate manner.

Auditors need to compare and contrast diverse sources of information, validate data integrity and accuracy and look for patterns and anomalies in data. The audit process needs to support assertions inherent in published financial statements such as completeness, uniqueness, accuracy, occurrence, valuation and presentation.

A purpose-built technology will have algorithms designed to perform these tests without having to program macros. Effective data analysis technologies for audit will have commands and functions that look for duplicates, detect gaps in numeric sequences, group transactions by type, numeric range, and age. The ability to filter vast amounts of data quickly and efficiently is also a key requirement. Advanced pattern detection techniques, such as digital analysis, are extremely helpful when seeking anomalies in data. In many audit situations, sampling is mandated or required. Both statistical sampling and monetary unit sampling are essential. When comparative analysis is required, the technology needs the ability to join data files (often from different sources and in different formats) and look for matched or unmatched records. For tasks requiring the comparison of data from numerous different tables, the ability to relate diverse datasets together is also necessary.

Since the audit process often involves retrospective analysis of vast amounts of data, an effective data analysis technology needs highly efficient read algorithms in order to process millions of records rapidly. These algorithms must be powerful and reliable in order to perform tasks either quickly in interactive data analysis or for sustained periods of time in lengthy and complex automated analysis.

Spreadsheets provide a view of data in columns and rows with the ability to enter text or images in cells or groups of cells. In this manner they are very adept and flexible – which in part explains their popularity. Spreadsheets also have a wide range of analytical capabilities for general data analysis, however, since they were not developed with the auditor’s needs in mind, they lack user-interface attributes that support the audit process. They do not feature the navigation capabilities required to organize data sources and associated data elements that provide critical context to the audit process.

Understandably, spreadsheets also lack most audit-specific analytics. It is with this point that spreadsheets demonstrate some of their strengths – which also highlight the risks associated with using spreadsheets in audit. Where specific analytics are missing from the standard suite, most spreadsheets enable the programming of macros to extend their analytical capabilities. For users who wish to create custom analysis and are not familiar with macros, they may be first inclined to chip away at the problem by creating complex formulas within the spreadsheet to get the results, for example: complex logical formulas, multiple cells referring to each other, combinations of linked cells/spreadsheets/formulas, lookups and more. These could increase the chance for errors and could make the spreadsheet more difficult to review. While this illustrates the flexibility and extensibility of spreadsheets, it also highlights a key weakness in how they are typically implemented. Macros are seldom designed, coded and tested according to sound software development practices. Even so, spreadsheets are rarely under version and configuration management or protected from editing. Even if a macro is effectively documented, there is no way of assuring its continued integrity.

More often than not, there is someone in the organization who is a “spreadsheet guru” who creates and shares macros without documenting what they do or testing the integrity of their output. The key point here is that unlike a professionally developed, audit-oriented data analysis product, those developing macros outside the oversight of the IT department do not carry the accountability or the responsibility of the results they generate. This inadvertently introduces significant risk into the organization. In the end, it is the individual auditor and the organization’s audit function that are at risk for providing opinions and making recommendations based on faulty results owing to poor software development practices.

Logging and Automation

One of the keys to improving audit performance and driving better results is the ability to automatically record what has been done and reliably repeat it in subsequent areas or in subsequent audits. It is for this reason that audit-specific data analysis technologies provide comprehensive audit trails. They also provide for reliable task automation, from accessing the data at source to verifying its validity to performing the detailed analysis and generating audit reports.

There are a number of attributes that constitute an effective audit trail. An effective audit trail is one that records all of the commands run by the application, status messages that provide insight into command execution, and any results generated by the actions of the user. This provides a number of critical artifacts for an effective audit, including a context for the audit findings.

Going back to the fraud-related example above, the chain of events that could have led to the identification of a fraudster may have been as follows:

1. Employee Master file is opened by the application as the primary file
2. Approved Vendor file is subsequently opened by the application as a secondary file
3. Join command is issued to identify matched records based on the Street Name field
4. Other fields specified in the command: Employee Name, Employee Number, Vendor Name, Vendor Number, City, State and Zip Code
5. Resultant file is created returning three different records showing matches in street, city, state and zip code
6. Resultant table is sorted, formatted and exported to be included in an audit report

This simple audit trail documented the steps taken to uncover these exceptions. The findings can now be explained, substantiated and defended where necessary.

The audit trail also provides a mechanism for peer or supervisory review. Review of audit steps is an important activity to ensure the accuracy and completeness of the audit process. This review demonstrates to audit management that opinions on the state of the organization's internal and external compliance are accurate and that the audit recommendations are sound.

What if the test conducted above neglected to specify the City, State or Zip Code of the matched records? It is quite possible that an employee lives on Broad Street in one city and that a vendor's address is on Broad Street in a completely different city.

In a completely different context, consider a case in which an auditor goes through a detailed and complex series of steps to identify cost and expense leakage in the accounts payable process. The work uncovers a large number of cases where contract discounts had not been recognized by a supplier, amounting to several hundred thousand dollars over a prolonged period. If the auditor neglects to manually document each step taken in this analysis using a spreadsheet, how can it be proven that these findings are based on a valid methodology? Are the steps repeatable? Only applications that consider the importance of the discovery process and its repeatability offer an acceptable solution for audit's needs.

A final benefit of an audit trail is the ability to recall previous results. An audit trail records not only the commands and functions used to identify exceptions and anomalies but also intermediary and final results. In this way, auditors may compare past findings with current findings to see if the recommendations have been acted upon, or if there is a substantive shift in the behavior of the business that may be an indicator of emerging risk.

For data analysis in general and for the audit process in particular, if meaningful results or insight are achieved through an analytical process, they are probably worth repeating again in the future. An effective audit technology enables simple and straightforward task automation. Effective technologies provide for a variety of ways to automate analytical tasks either through “task recorder” type functionality or through the selection of commands recorded in the audit trail.

Ideally, the scripting language should be simple enough for a non-technical individual to understand and edit without breaking the integrity of the script. It is through straightforward task automation that auditors themselves may create batteries of tests to streamline the overall audit process and lead towards a continuous auditing framework.

Suitability for enterprise-class deployment

The audit function is one that spans the entire organization. It is not limited to activities controlled by a single application or recorded in a single database. Audit needs to gain insight into the ongoing financial performance and operations of the organization in order to assess risk, opine on the efficacy, suitability and adherence to controls and regulations, and make valuable and insightful recommendations to the audit committee. To deliver on this tall order, audit needs visibility into all aspects of the organization to evaluate the performance of the organization through independent analysis of huge volumes of transactions recorded in a myriad of databases, data files, spreadsheets and reports. An effective technology to support this must be able to read and interpret vast amounts of data in any number of ways from a single and intuitive user interface.

Data must be compared and contrasted, joined and related, stratified and aged into different categories and buckets. The ability to drag columns around views and apply data filters quickly and sort data with immunity from altering the source data is essential. Furthermore, extracting data and distributing it to insecure and unmanaged locations and applications must not compromise the organization’s control and security of the data.

An effective solution will allow data to be analyzed in place – on the servers and in the databases in which it resides. Ideally, this will be from a single intuitive user interface where users move seamlessly between all types of data sources and formats. The complexities of server operating systems and configurations should remain hidden from the user. Whether the data resides on networked PCs, mainframes, or servers, an effective audit technology will provide a single point-of-view into your organization’s data. This allows the auditor the capability to move beyond silos of data. These could be silos of data as exists within business applications, enterprise resource planning (ERP), customer relationship management (CRM), supply chain management (SCM), human resource management systems, or function-specific applications and legacy systems. These are multiple, and often incompatible, data sources. An enterprise-class data analysis application provides visibility and insight into the integrity of business transactions and is not constrained by its ability to be converted into a format that a spreadsheet is able to analyze.

COMPANY BEHIND THE PRODUCT

When selecting an appropriate data analysis product for audit, one should consider the software publisher and the type of supporting services provided to end-users. When deciding on which technology to deploy for the audit process, one should seek a “whole product” provider.

A whole product provider is one that supplies the software product, supporting documentation, Help Desk services, training offerings, implementation services and consulting services. In order to get maximum value from your software investment, support services that complete a whole product offering should be focused on the auditor and the audit process. This means the Help Desk understands the tasks that auditors are executing with their software tools and the audit objectives they wish to realize. This applies to all steps necessary in the data analysis tasks of audit: data access, data validation and cleansing, use of commands, expressions, filters and task automation. Similarly, training courses provided on the software should be audit-oriented and aligned with the high-value tasks involved with audit. Implementation and consulting services should be focused on accelerating audit productivity and efficiency. The underlying expertise that a whole product provider must have, like the technology itself, is domain-specific.

This approach is in stark contrast to the services offered by spreadsheet providers.

While spreadsheet providers offer excellent support services of their own, their services, like their products, are not specific to data analysis in audit.

CONCLUSION

Spreadsheets are one of the most popular PC applications in business use today. The usefulness, flexibility and power of spreadsheets are undeniable, making them an indispensable tool for a broad range of applications. Spreadsheets fill some critical gaps in financially material business processes today and in the foreseeable future. The audit function within an organization is charged with the vital task of systematically and diligently evaluating and improving the risk management, control and governance processes. Audit requirements and responsibilities are such that the risks inherent in spreadsheet use for data analysis are far too great. Audit requires select, purpose-built technologies to inoculate organizations against a variety of risks and threats – some of which may even be posed by the use of spreadsheets.



ACL Headquarters

T +1 604 669 4225

F +1 604 669 3557

**■ acl.com
info@acl.com**

■ ■ ■ COMPANY OVERVIEW

ACL is the leading global provider of Business Assurance Analytics to financial executives, compliance professionals, and auditors. Combining market-leading data analytics software and professional services expertise, ACL solutions give organizations confidence in the accuracy and integrity of transactions and the effectiveness of internal controls underlying increasingly complex business operations. Since 1987, ACL's proven technology has enabled financial decision-makers to assure controls compliance, reduce risk, detect fraud, enhance profitability, and achieve fast payback.

ACL delivers its solutions in more than 130 countries through a global network of ACL offices and channel partners. Our customers include 70 percent of the Fortune 500 companies and over two-thirds of the Global 500, as well as hundreds of national, state, and local governments, and the Big Four public accounting firms.