|  |  |
| --- | --- |
| **Ivey Publishing logo** | Bond University logo |

9B18B013

an introductory note on BIG DATA AND DATA ANALYTICS FOR ACCOUNTANTS AND AUDITORS

Molly Parker wrote this note under the supervision of Dr. Adrian Gepp solely to provide material for class discussion. The authors do not intend to illustrate either effective or ineffective handling of a managerial situation. The authors may have disguised certain names and other identifying information to protect confidentiality.

This publication may not be transmitted, photocopied, digitized, or otherwise reproduced in any form or by any means without the permission of the copyright holder. Reproduction of this material is not covered under authorization by any reproduction rights organization. To order copies or request permission to reproduce materials, contact Ivey Publishing, Ivey Business School, Western University, London, Ontario, Canada, N6G 0N1; (t) 519.661.3208; (e) cases@ivey.ca; www.iveypublishing.ca. Our goal is to publish materials of the highest quality; submit any errata to publishcases@ivey.ca.

Copyright © 2018, Ivey Business School Foundation Version: 2022-09-28

This note provides an overview of big data and data analytics, and how they can be effectively used in accounting and auditing practices. Historically, data analytics have not been used in the professions of accounting and auditing to the extent that they have been in other industries. However, an understanding of big data and the relevant analysis techniques is becoming almost mandatory for accountants and auditors so that they can provide greater business insights for their businesses and clients.

The note begins with an overview of why the field of big data is important for accountants and auditors. The reader will then learn about the related, yet separate, concepts of big data, data analytics, and data analytics tools from an accounting perspective. The three stages of the data analytics process are also explained, followed by a brief introduction to ACL Analytics, a widely used data analytics tool available to assist auditors. The introductory overview brings previously discussed concepts into context and leads to a concluding discussion about the growth and potential future applications of big data and data analytics in the fields of accounting and auditing. A detailed tutorial on the use of ACL Analytics is offered in the Appendix; it provides the information the reader needs to apply the concepts covered in this note in practice. The tutorial also enables the reader to gain practical skills in the basics of ACL Analytics.

Introduction

Big data and data analytics are becoming increasingly relevant in modern times. As noted in *Harvard Magazine*, “Data now stream from daily life: from phones and credit cards and televisions and computers; from the infrastructure of cities; from sensor-equipped trains, buses, planes, bridges, and factories.”[[1]](#footnote-2) The massive data sets that are available, combined with the ever-developing data analytics tools and techniques, offer businesses the opportunities to add value and generate greater insights. However, the professions of accounting and auditing have not yet capitalized on these opportunities to the same extent as have other industries. For accounting and auditing to remain relevant in the modern business world, it is essential that the industry as a whole increases its use of relevant data analytics techniques and tools. For this increased prevalence to materialize, individual accountants and auditors must develop their own personal understanding of big data and data analytics so they may evolve in time with the industry.

IMPORTANCE OF BIG DATA IN ACCOUNTING AND AUDITING

In 2018, with information being readily accessible and created at an unprecedented pace, it is vital that those working in professional services understand how to use these data to their advantage. Without this understanding, information overload is likely and important data may be ignored. Such an outcome will result in a loss of competitive position to those entities that effectively use the data.

It is important to clarify here the difference between big data and data analytics. The term “big data” is intuitive, in essence referring to extremely large data sets, while “data analytics” refers to the process of transforming and analyzing these large data sets to produce information that can be effectively used. Most businesses currently possess big data, but this ownership alone is not enough; the analytics are vital to make sense of and use the information to improve decision-making and business outcomes.

Many benefits accrue to accountants and auditors who understand how to use big data analytical tools and software. Most importantly, data analytics offers them a competitive advantage in adding value to their business activities, whether working as an auditor or as an accountant for a firm in any industry. Some experts have referred to the current lack of understanding regarding data analytics as the leading challenge for firms adopting such processes—a challenge greater than dealing with data quality or governance.[[2]](#footnote-3)

Amazon is a company that effectively uses data analytics to its advantage. Its selling recommendation algorithms, based on information about customers’ online shopping habits, have been hugely successful and generate a substantial percentage of Amazon’s revenue. However, the professions of accounting and auditing, in general, have yet to capitalize, to the extent that Amazon has, on the opportunities that data analytics offers.

Accounting and finance are already using data analytics to develop automated models to predict the future financial distress (or failure) of firms, to detect financial fraud, and to make stock market predictions, but auditing is lagging behind, not yet using modern data analytics on a substantial scale.[[3]](#footnote-4) Research into the use of data analytics by accounting and auditing firms is also limited. However, the professions of accounting and auditing are beginning to realize the valuable business insights that are being missed. The Big Four accounting firms, for example, all have pages on their websites that promote their use of big data—evidence that they are beginning to emphasize their use of data analytics within their businesses.[[4]](#footnote-5) Private communication between the authors and partners and staff at such firms also supports this increase in emphasis.

Data analytics enables audit firms to easily provide clients with new and valuable insights from their data. A technique known as “data bridging” can offer insight and add value to an audit by linking traditional data sets with external, non-traditional data. For example, photos, videos, and global positioning system locations could all be used to verify transactions, rather than relying on only more traditional receipts or invoices.[[5]](#footnote-6) Another technique links audited sales data with demographic and socio-economic information, such as census data. In addition to verifying an increase in sales over the previous year, auditors could use this information to determine, for example, the common demographic and socio-economic factors in regions where sales are underperforming. This information is a valuable additional insight to share with clients.

Whole population analysis is yet another data analytics technique that can add value to businesses. Traditionally, transaction tests in audits have been done by manually testing a small sample of transactions and inferring results about the entire population of data. However, whole population analysis, which performs tests on the entire data set, enables auditors to quickly and efficiently identify all transactions that meet or exceed a certain risk profile, thus allowing the auditors to easily identify and focus on more important areas of the audit. The accuracy that whole population analysis provides gives audits more weight than was previously possible, and offers both the audit firm and the client a competitive advantage in their respective business areas. Extra services such as these are becoming almost expected.[[6]](#footnote-7)

The audit process itself also becomes more efficient when both the auditor’s and the client’s accountants understand big data and the relevant analytics techniques. The first efficiency gain is improved communication between the client’s accountants and the audit team. Quicker and more effective conversations are possible when both parties understand the data analytics process being used. Clients are able to easily answer the audit team’s questions about the data and data sources, and can efficiently extract the requested data from their systems. The audit team benefits by being able to directly question and provide answers to the client relating to basic data issues without needing to first go through a data analytics team. An added benefit is that the data analytics team, which is a specialized resource, need not be involved in all the basic steps, which would reduce their efficiency without any offsetting benefits. This approach requires, however, that all auditors and accountants have at least a basic understanding of common data analytics tools, techniques, and terminology to stay current with their growing use in the industry and, at the very least, be able to communicate effectively with data analytics specialists.

It is essential to mention that auditors must continue to safeguard their independence. However, sharing insights garnered from performing data analytics that are relevant to the audit do not generally pose considerable threats to independence. For example, if an auditor determines that a large amount of inventory would take three years to sell, then it would be appropriate to share that insight with the client, as well as request adjustments to inventory valuation as necessary. However, performing data analytics to provide a client with helpful information that is not relevant to the audit can pose greater threats to independence. Auditors must particularly watch for advocacy threats (i.e., advocating matters for the client) and management participation threats (i.e., making management decisions). Importantly, doing so does not undermine the relevance and importance of data analytics within audit engagements; rather, it reaffirms the ongoing need to be vigilant in complying with independence standards, particularly when performing analysis that is not directly relevant to the audit engagement.

The International Standards on Auditing (ISAs) do not prohibit the use of data analytics, but their lack of reference to them could be a barrier to their widespread adoption.[[7]](#footnote-8) Nonetheless, while the standards are not changing as fast as the rise of data analytics, the International Auditing and Assurance Standards Board (IAASB) acknowledges the value of data analytics and that data analytics will be an important part of many standards in the future. The IAASB has indicated that data analytics will play a role in the considered changes to ISA 315 (Revised), “Identifying and Assessing the Risks of Material Misstatement through Understanding the Entity and Its Environment” and ISA 540, “Auditing Accounting Estimates, including Fair Value Accounting Estimates and Related Disclosures*.*”[[8]](#footnote-9) The IAASB is also planning a preliminary analysis of how data analytics might be incorporated in ISA 520, “Analytical Procedures,” and has noted that data analytics may also affect ISAs 240, 320, 330, 500, and 530 in the future.[[9]](#footnote-10) These plans are further evidence that auditors of the future will need to be proficient in data analytics.

Yet despite their benefits and increasing acceptance, data analytics tools are still significantly underused in auditing. Computer-assisted auditing techniques (CAATs)—any technology used in the auditing process—have been used for many years, but these tools are becoming inadequate to handle the increasingly large data sets that clients provide.[[10]](#footnote-11) In 2018, CAATs that are more advanced than the traditional tools of spreadsheets, word processors, and simple databases have become available; these can be referred to as data analytics tools. While big data and data analytics are being used at some level by auditors, there is still limited awareness of these techniques and what they are able to do, among both auditors and clients. Both audit and client firms will often exclusively depend on specialized data analytics teams to perform data-intensive tasks. However, simply relying on a small team of data specialists is insufficient when attempting to use such techniques on a large scale and for new purposes. The data analytics teams need to be fully integrated with the rest of the audit process so that opportunities to use new techniques can be identified.

CONCEPTS OF BIG DATA, DATA ANALYTICS, AND DATA ANALYTICS TOOLS

The first step to effectively providing data analytics services to all clients is to be aware of and understand what data analytics can offer audit firms and their clients. The following section provides an introduction to these concepts, specifically tailored for auditors and accountants.

Big Data

The term “big data” refers to structured or unstructured data sets—structured referring to data organized in a database or a database-like format (such as spreadsheets), which can be filtered using straightforward search engines or algorithms. Unstructured data, on the other hand, are data that are not in a strictly organized format and are generally easier for humans rather than computers to use and analyze. An email inbox would be an example of unstructured data because there is no way to precisely sort the information by subject or content. Video and audio data are other examples. Overall, the amount of unstructured data is increasing at a rapid rate.

Big data sets can also be defined by the four Vs: huge Volume, high Velocity, wide Variety, and uncertain Veracity. Huge volume refers to data sets that are too large to be stored or analyzed using traditional analytics tools, although being “too large” is relative for firms of different sizes. For example, a data set that is considered big by a mid-level accounting firm would not be considered big by NASA (National Aeronautics and Space Administration). High velocity refers to the speed at which new data become available. With the introduction of real-time information and continuous accounting systems, the velocity of data is increasingly growing, and can be almost instantaneous in some cases. Wide variety refers to different and diverse formats of structured and unstructured data, such as audio, image, or video, as well as the more traditional quantitative or text formats. The final V, uncertain veracity, refers to the quality and relevance of data over time. The veracity of data sets can change dramatically: consider, for example, the reliability of data provided through a third-party system or mobile application (app) where the owning company may be acquired one or more times.

Data Analytics

In the practices of accounting and auditing, big data are becoming increasingly prevalent as businesses discover new, different, and more effective ways of reporting their statistics and financial information. This trend is evidenced by the American Accounting Association having introduced an annual conference entitled “Accounting IS Big Data.” However, these data are unusable unless accountants and auditors know and understand the analysis and its application to business practices.

Data analytics can be described through the process of extracting, transforming, and loading data, often referred to as ETL.

* *Extracting* refers to transferring raw data from a database or accounting system. However, big data are beginning to pose challenges to the extraction part of the data analytics process, as a result of both increasing velocity and variety of data. High velocity data require automated extractions because the data are arriving too fast to be handled manually. Further, data provided in a variety of formats such as audio or video often need to be extracted from non-traditional sources. Current tools for these types of data extraction are not as developed as those for the extraction of traditional data, but they are quickly improving.
* *Transforming* is the most substantial part of the ETL process; it refers to altering the raw data into a format that can be easily used. The process can range from simple transformations of financial statements to linking numerous big data sets together and complex multivariate modelling. Data transformation can be performed using a variety of tools and techniques, an example of which is discussed in more detail later.
* *Loading*, the final part of the process, refers to the act of transferring or uploading the transformed data into the database or accounting system, where the data can then be used effectively.

In recent years, data analytics techniques have begun to automate tasks and procedures traditionally performed manually by accountants and auditors, such as posting and collecting accounts receivable.[[11]](#footnote-12) These automated tasks may lead some to assume that the “big data revolution” could result in accounting professionals becoming obsolete. However, the ETL process and the data analytics tools still require human experts for efficient use and to truly achieve the potential value from the analysis.[[12]](#footnote-13)

The types of tasks that have been automated, or will be automated in the near future, are those that are quite simple and have traditionally been performed by employees in entry-level positions.[[13]](#footnote-14) Once these tasks have been automated, humans are still required to continue to adapt the automatic systems, as well as to verify that the systems continue to operate as they should. The main challenge that will be posed to accountants and auditors is to develop skills in the more intricate and tailored data-related tasks where the high-quality insights will be found.

Data Analytics Tools

Numerous tools and software can be used to assist professionals in performing data analytics. These tools, while varied in type and method, generally have one final aim: to assist in transforming big data sets from a format that is too large to comprehend into a format that can be easily used to realize business insights.

Many data analytics programs are designed to transform data sets into a general and consistent format, which is extremely useful when using those programs and analytics on a large scale. From this generalized data format, it is easy to create and use data visualization tools such as graphs, tables, and the increasingly popular dashboards. The consistent format also makes it easier to reuse code on future data sets, which increases efficiency and enables experts to focus their time and energy on tailoring the output for each client. In addition to statements on their corporate websites, anecdotal evidence from discussions with multiple audit partners in Australia reveals that many firms have already begun to use data analytics tools in this manner.

ACL Analytics (ACL), developed and sold by ACL Services Ltd., is a widely used data analytics tool in the audit profession. A coding and analytical software package that uses the audit command language, ACL can store and analyze large quantities of data. ACL’s main competitor is CaseWare IDEA, although ACL is more commonly used since it is older and more established.[[14]](#footnote-15)

The next section of this note contains a more detailed introduction to ACL as an example data analytics tool, discussing what it is, what it can do, and the basics of how to use the program. Other software programs also perform work very similarly to ACL, and, as a result, the introductory lessons taught here can be extrapolated for use with other coding and analytical software (also called script-based tools), such as IDEA, SAS, or SQL. Whether or not the reader intends to perform the analysis, these lessons will improve the reader’s understanding of both the data analytics process and how it can be used by audit firms.

ACL Analytics

ACL uses three types of files: tables, scripts, and workspaces. *Tables* are files that contain imported data sets, whereas *scripts* and *workspaces* are files in which code can be written. The code within scripts and workspaces is used to transform existing tables, as well as create new ones. The step-by-step, introductory ACL tutorial in the Appendix uses only tables and scripts.

The tutorial works through a basic example of a script to combine two tables in ACL—a good example of simple data transformation. The tutorial includes written sections of code and screenshots of the program interface to demonstrate the output that is being generated. The two tables are trial balances, one from the beginning-of-period and one from the end-of-period; for example, January 1, 2017, to December 31, 2017. The data was created for the purpose of the tutorial and are not real-world financial data.

Once the trial balances (TBs) are combined into one table, ACL is able to quickly reconcile those data against the general ledger (GL) journal entries for the same period. This reconciliation allows the user to check for risky transactions, such as dates outside the period, unauthorized users, and unusual transaction times or patterns. In this tutorial, only the TB tables and scripts are worked through in detail, but similar processes and functions can be used with GL data.

The automation of the data transformation process that ACL or other similar software offers enables auditors to add value to their business and for their clients. This added value is, in part, due to efficiency: once scripts are set up, they can be reused, with minor changes, for numerous data sets. This ability helps to automate parts of the extract, transform, and load process, which is incredibly time-efficient. In addition, the ability to test whole data sets rather than inferring results from a sample is incredibly valuable for an audit.[[15]](#footnote-16)

The lessons from this overview of ACL and the tutorial in the Appendix are applicable in numerous areas because most script-based data analytics programs are similar. Having a basic understanding of how functions and commands work is helpful when using any sort of coding software. The detailed tutorial also helps to visualize what coding analytics software looks like, which is useful in overcoming the first impression that these programs are intimidating and complicated. The specific and detailed example of data transformation in the tutorial also places the theory from previous sections of this note into context.

While the tutorial is introductory, programs such as ACL offer great potential in their higher-level analytics and outputs. However, not all processes can be automated; human experts are required to customize the code to specific engagements, troubleshoot errors, and, most importantly, interpret the results. Often reconciliations of the GL to the TB will reveal large numbers of differences, usually a result of issues with the source data or a mistyped line of script. Troubleshooting these errors is unlikely to be automated in the near future, but troubleshooting does require human experts to have knowledge of data analytics.

DISCUSSION AND CONCLUSION

The fields of accounting and auditing have substantial potential for efficiency gains through the use of data analytics. The ACL tutorial in the Appendix provides just one example of how traditional accounting and auditing processes are becoming quicker and more efficient. However, individual accountants and auditors, and the profession as a whole, still have much to do to develop the skills and acquire the understanding necessary to access the full potential that data analytics offer.

A survey on ACL use asked auditors to respond to a series of statements with a number between −3 and 3, with −3 being “strongly disagree” and 3 being “strongly agree.”[[16]](#footnote-17) The full results of the survey can be found in Exhibit 1. The statement with the most agreement (average of 2.29) was “Using ACL improves overall audit effectiveness,” however, the statement with the least agreement (average of −0.54) was “I have never encountered any significant problems in using ACL.” These results indicate that while ACL is generally accepted as being able to increase the value of audits, users within auditing firms often face issues when using the software. Although this survey, done in 2003, is slightly dated, it highlights the need for appropriate training for auditors and accountants on the relevant data analytics techniques, tools, and software (such as ACL). Without sufficient education, auditors and accountants will be unable to realize the full potential that such techniques and tools can offer.[[17]](#footnote-18) This education needs to come through universities and be provided as on-the-job training, for both new and established professionals.

The growth of big data and data analytics could be seen as a threat to the accounting profession, but it is more constructive to view the emerging use of data as an opportunity. Experts suggest that in order for firms to effectively use big data and data analytics, they should employ people who understand not only the data analytics but also the business.[[18]](#footnote-19) If accountants and auditors develop the understanding and skills related to data analytics, they would be perfectly poised to fill these new positions because of their pre-existing understanding of the business activities and strategies.

The wider profession of auditing and accounting has been relatively slow and lagging behind other industries on the uptake of data analytics techniques. However, many experts believe that use of data analytics will increase rapidly once the potential is fully realized and appreciated.[[19]](#footnote-20) In fact, audit clients have already started to ask more about the use of analytics and to even expect it.[[20]](#footnote-21) It is, therefore, more important than ever for individual professionals working in these areas to understand and apply big data and data analytics.

While this overview of ACL and the included tutorial are introductory, an appreciation of how the tools and techniques work at a basic level helps auditors and accountants to gain greater insights from high-level commentary and discussions on the topic. Increased education and training, for both accounting students and current professionals, would further enhance their knowledge of and ability to use big data, data analytics, and specific processes and tools, such as ACL. In this way, accountants and auditors can be well equipped to capitalize on the opportunity that the growth of data analytics provides to both individual professionals and the industry as a whole.

Exhibit 1: Survey of ACL Use among Auditors

Survey Results, by Average Response

|  |  |
| --- | --- |
| **Statement** | **Average Response** |
| Using ACL improves overall audit effectiveness | 2.29 |
| I am able to complete audit procedures more efficiently using ACL than I could without ACL | 2.00 |
| I am generally receptive to change | 2.00 |
| I would be interested in participating in more ACL training if it were available | 1.98 |
| Instructor-led training from ACL is worthwhile | 1.85 |
| I consider myself technologically competent | 1.78 |
| I feel confident in designing and interpreting basic ACL commands once the input file has been defined | 1.76 |
| I would make use of a “command library” consisting of ACL commands that have been beneficial in other audits if such a library were available | 1.72 |
| My documentation of use of ACL is adequate to allow future auditors to understand what I did | 1.69 |
| I have received good support from Information Systems Audit personnel within the office | 1.63 |
| The input file definition process is beneficial toward obtaining a better understanding of client data | 1.63 |
| I tend to do less work on representative samples and more work on high-risk samples than I would have if I did not use ACL | 1.51 |
| Using ACL increases the likelihood of audit findings | 1.37 |
| Much of my training on ACL comes from colleagues | 1.33 |
| I have encountered system problems (e.g., lockouts, dial-up, etc.) that have impaired efficiency | 1.24 |
| I feel confident with the process of creating input file definitions | 1.24 |
| Using ACL allows me to gain a clearer understanding of a matter before referring it to the investigative division | 1.21 |
| I would generally be able to identify the basic contents of most ACL files created on my most recent audit by looking at the names of the files | 1.18 |
| I feel confident in designing moderately complex commands (such as “join” commands, etc.) | 1.12 |
| Other auditors would be able to gain a rough idea of the content of my ACL files by looking at the names of those files | 1.10 |
| I feel well trained in the use of ACL | 1.00 |
| I believe my colleagues are well trained in the use of ACL | 0.91 |
| Most other auditors’ documentation of use of ACL is adequate to allow understanding of what was done | 0.80 |
| Using ACL increases the likelihood of referrals to the investigative division | 0.78 |
| Efficiency gains from using ACL usually occur in the first year of use | −0.01 |
| I have encountered significant ACL-related problems on my most recent audit | −0.48 |
| I have never encountered any significant problems in using ACL | −0.54 |

Exhibit 1: Continued

Survey Results, by Statement Category

|  |  |
| --- | --- |
| **Statement** | **Average Response** |
| **Potential Benefits** |  |
| Using ACL improves overall audit effectiveness | 2.29 |
| I am able to complete audit procedures more efficiently using ACL than I could without ACL | 2.00 |
| I tend to do less work on representative samples and more work on high-risk samples than I would if I did not use ACL | 1.51 |
| Using ACL increases the likelihood of audit findings | 1.37 |
| Using ACL allows me to gain a clearer understanding of a matter before referring it to the investigative division | 1.21 |
| Using ACL increases the likelihood of referrals to the investigative division | 0.78 |
| Efficiency gains from using ACL usually occur in the first year of use | −0.01 |
| **Potential Problems** |  |
| I have encountered system problems (e.g., lockouts, dial-up, etc.) that have impaired efficiency | 1.24 |
| I have encountered significant ACL-related problems on my most recent audit | −0.48 |
| I have never encountered any significant problems in using ACL | −0.54 |
| **Technical Skills** |  |
| I feel confident in designing and interpreting basic ACL commands once the input file has been defined | 1.76 |
| The input file definition process is beneficial toward obtaining a better understanding of client data | 1.63 |
| I feel confident with the process of creating input file definitions | 1.24 |
| I feel confident in designing moderately complex commands (such as “join” commands, etc.) | 1.12 |
| **Training and Support** |  |
| I would be interested in participating in more ACL training if it were available | 1.98 |
| Instructor-led training from ACL is worthwhile | 1.85 |
| I would make use of a “command library” consisting of ACL commands that have been beneficial in other audits if such a library were available | 1.72 |
| I have received good support from Information Systems Audit personnel within the office | 1.63 |
| Much of my training on ACL comes from colleagues | 1.33 |
| I feel well trained in the use of ACL | 1.00 |
| I believe my colleagues are well trained in the use of ACL | 0.91 |
| **Documentation Practices** |  |
| My documentation of use of ACL is adequate to allow future auditors to understand what I did | 1.69 |
| I would generally be able to identify the basic contents of most ACL files created on my most recent audit by looking at the names of the files | 1.18 |
| Other auditors would be able to gain a rough idea of the content of my ACL files by looking at the names of those files | 1.10 |
| Most other auditors’ documentation of use of ACL is adequate to allow understanding of what was done | 0.80 |
| **Personal Characteristics** |  |
| I am generally receptive to change | 2.00 |
| I consider myself technologically competent | 1.78 |

Exhibit 1: Continued

* Participants were asked to rate a series of statements on a scale from −3 to 3, with −3 being “strongly disagree” and 3 being “strongly agree.”
* An email with a link to an online survey was sent to the directors of training at each of the legislative audit offices in the United States. The directors then forwarded the link to 350 auditors, 93 of whom visited the web page and 90 of whom completed the survey.
* The 90 auditors who completed the survey covered 14 U.S. states.
* Thirteen survey participants had 0–2 years of experience in the audit profession, 23 had 3–5 years, 22 had 6–10 years, and 31 had greater than 10 years.\*
* Because this survey was voluntary, it is possible that those who responded were more receptive to audit technology and data analysis than auditors in general.

\* In the original source, the breakdown of the 90 participants by experience level accounted for 89 participants.

Source: Robert Braun and Harold Davis, “Computer-Assisted Audit Tools and Techniques: Analysis and Perspectives,” *Managerial Auditing Journal* 18, no. 9 (2003): 725–731.

Appendix: ACL Analytics Tutorial

Imagine you are an auditor for a Big 4 firm, and your current engagement is Integrated Components Worldwide (ICW), a manufacturer of electrical components used in computers, drones, and other devices. You have requested detailed listings for one of ICW’s entities for the period January 1, 2017, to December 31, 2017. ICW has provided you with the opening and closing trial balance (TB) files and the general ledger (GL) for this period (See Ivey products 7B18B013A and 7B18B013B). These files were provided to you unformatted, in tab-delimited text format. Excel versions are also available (See Ivey products 7B18B013AX and 7B18B013BX). You wish to use these files to test for any unusual transactions. This tutorial details the first steps in this process.

The data used in this tutorial (TB\_Opening and TB\_Closing) have been provided so the reader can follow the steps shown and gain a deeper understanding by performing the tasks and not simply reading. It is highly recommended that readers replicate this tutorial themselves if they have access to the required ACL software. Coding can appear daunting to those with little or no experience; however, to best understand and appreciate the capabilities of coding software, it is important to see how coding is done and how best to perform it first-hand.

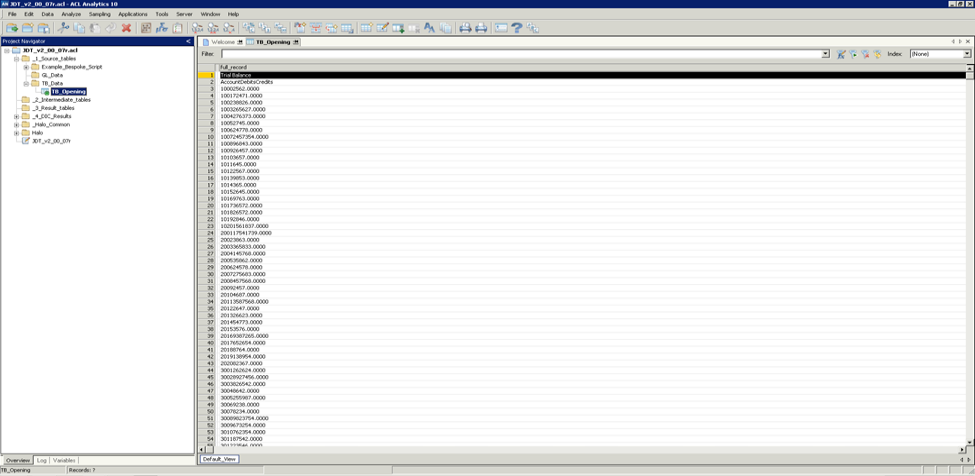
For easy reference, the final section of this appendix includes a list of all functions and commands used in this tutorial, as well as several other useful functions worth exploring in ACL.

1. Script Set-up and Opening TB

We start with opening and closing TBs for a period, which we want to combine into one table. The source data are in the format of tab-delimited text, meaning the columns are separated by tabs (represented by “>” in ACL).

When the data are initially imported into ACL, the data are not yet split into columns (see Figure 1: Opening TB Table). Consequently, the first step in writing a script to combine the opening and closing TBs is to separate the columns.

Figure 1: Opening TB Table



Set Folder /\_1\_Source\_tables/TB\_Data

Open TB\_Opening

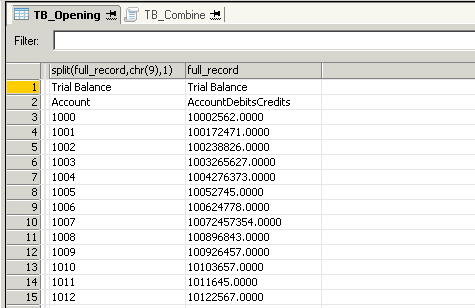
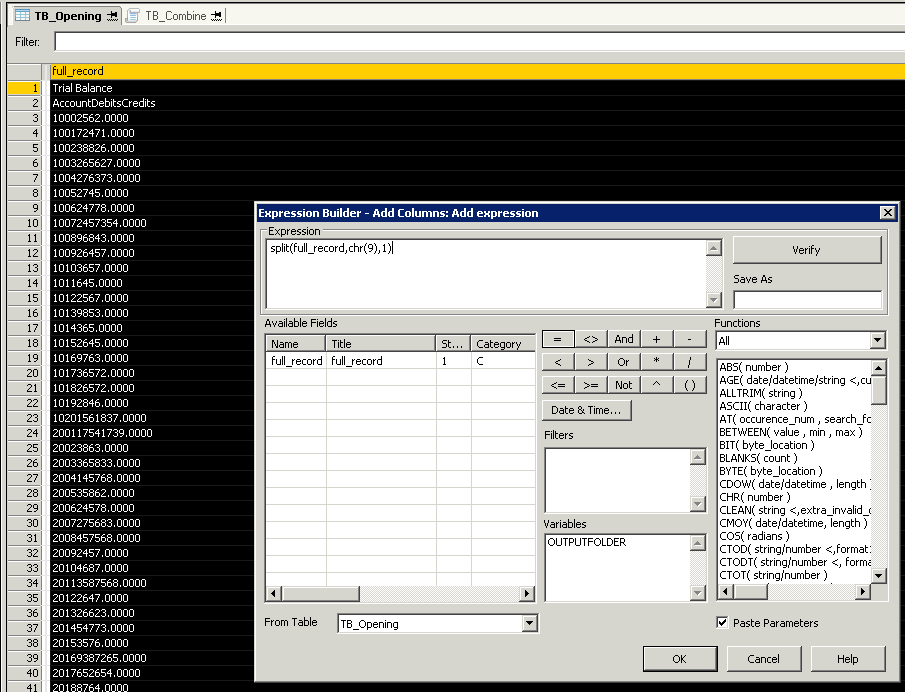
Extract fields to temp\_TB

Appendix (continued)

This is a good way to set up the script. The first line sets which folder the new combined table will be saved to (as seen in the menu on the left-hand side of the screen of Figure 1). We then instruct ACL to open the table TB\_Opening and extract the fields to a new table, which we are calling temp\_TB. This new table is temporary (similar to a draft), which we will then adjust slightly to create the final TB table later in the process.

1. Separating Columns

Figure 2: Separating Columns



Appendix (continued)

In Figure 2: Separating Columns, we add a column to the TB\_Opening table. The purpose is to check that the correct function is used in the TB\_Combine script. As can be seen, the function SPLIT(full\_record, chr(9), 1) separated the Account column from the rest of the data. The Split function follows the format of SPLIT(string, separator, segment, <text\_qualifier>). Here, string refers to the data we want to split, separator refers to the character we are separating by, segment refers to which section of separated data we want, and <text\_qualifier> is an optional argument that tells ACL what character is used to signify text.

Using chr(9) as the separator argument lets ACL recognize that it needs to separate by tab; we could also use a tab symbol in quotation marks (“>”) for the same result. We then put this entire split function inside another function—SUBSTR() or SUB().

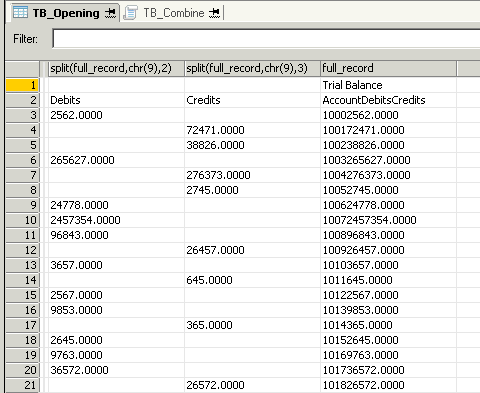
SUB(SPLIT(full\_record, chr(9), 1), 1, 4) as 'Account\_Number'

Following the format of SUB(string, start, length), this function simply tells ACL that the column containing the account number should start with the first character and end with the fourth character. This function removes any extra characters (e.g., spaces) that may have been in the source data. While this function may not initially seem important, it is an issue that commonly occurs in real-world data. This function is also helpful when only part of the information that is in a particular column is useful. An example is a situation where account numbers begin with an entity code; to separate a three-character entity code from the rest of the account number, the SUB() function could be used with 1 and 3 as the start and length arguments, respectively.

1. Debits and Credits

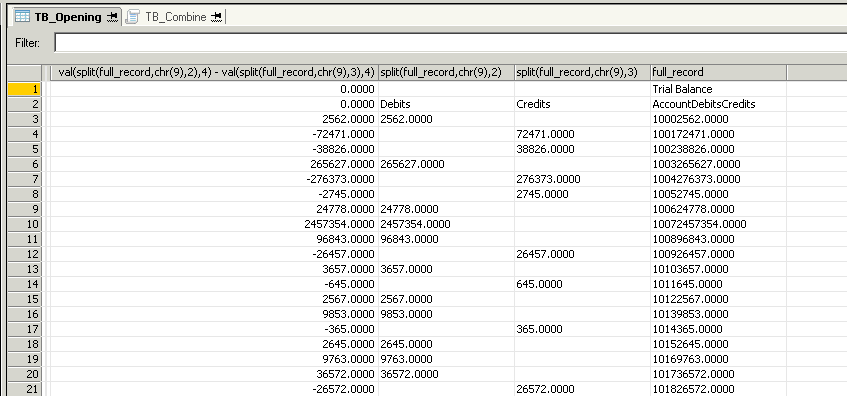
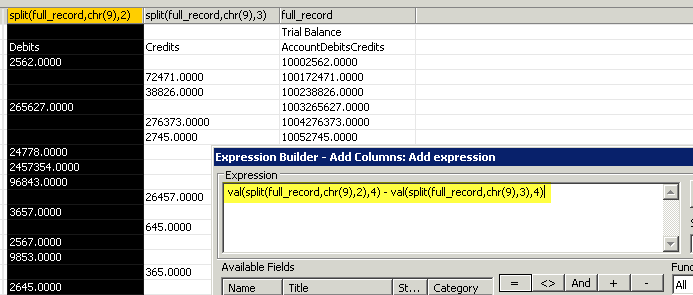
We next want to return to the TB\_Opening table and split out the debit and credit columns using SPLIT(full\_record, chr(9), 2) and SPLIT(full\_record, chr(9), 3) respectively (see Figure 3: Debits and Credits). However, we want to combine these debits and credits into one column, which is the simplest way to reconcile the GL to the TB further along in the process. To do this, we first must instruct ACL to recognize the data as numerical values rather than as text, using the VALUE(string, decimals) function. This can also be written as VAL(). We then want to subtract the credits (column 3) from the debits (column 2) using the function seen in Figure 4: Value Function. This action leads to a combined column for both debits and credits, showing debits as positive values and credits as negative values.

Figure 3: Debits and Credits



Appendix (continued)

Figure 4: Value Function



We can see here that this function is correct, and as such we can copy it into the script as Opening\_Amount. We must then hardcode the closing amount as 0, again using the VAL() function. This step is needed because the table we are currently extracting from includes only opening TB data; we will factor in closing TB data when we extract from the closing TB table.

VAL(SPLIT(full\_record, chr(9), 2), 4)

- VAL(SPLIT(full\_record, chr(9), 3), 4) as 'Opening\_Amount'

VAL("0",4) as 'Closing\_Amount'

Now that we have split out all three columns from the source data, we are required to add one more column: the name of the business. This step identifies different data within the same project, which becomes important when the client has numerous entities. Because the name of the business is not included in the source data, we can hardcode it using the SUB() function.

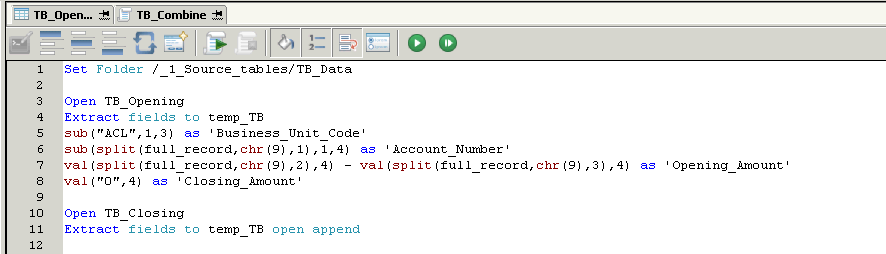
SUB(“ACL”, 1, 3) as ‘Business\_Unit\_Code’

1. Closing TB

Now that we have written all of the necessary code for the opening TB, we must append the closing TB to the opening TB in the temp\_TB table that we are creating (see Figure 5: Appending TB Closing Table). The open append command tells ACL to put the data from the TB\_Closing table directly below that of TB\_Opening, all in the one table.

Appendix (continued)

Figure 5: Appending TB Closing Table



Because the opening and closing TBs are in the same format, appending the closing TB is as simple as copying lines 5 to 8 in the script and pasting them at line 12. However, we must swap the two value functions so that the opening amount is hardcoded as 0 and the closing amount is populated with data.

Open TB\_Closing

Extract fields to temp\_TB open append

SUB("ACL", 1, 3) as 'Business\_Unit\_Code'

SUB(SPLIT(full\_record, chr(9), 1), 1, 4) as 'Account\_Number'

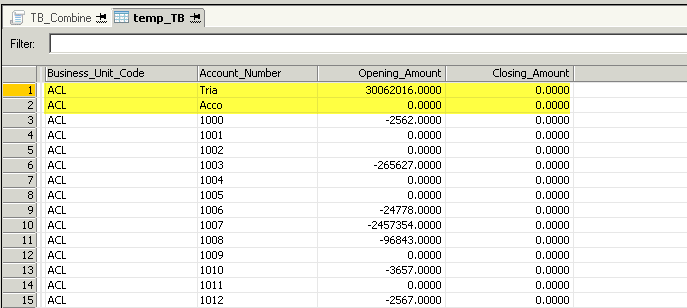
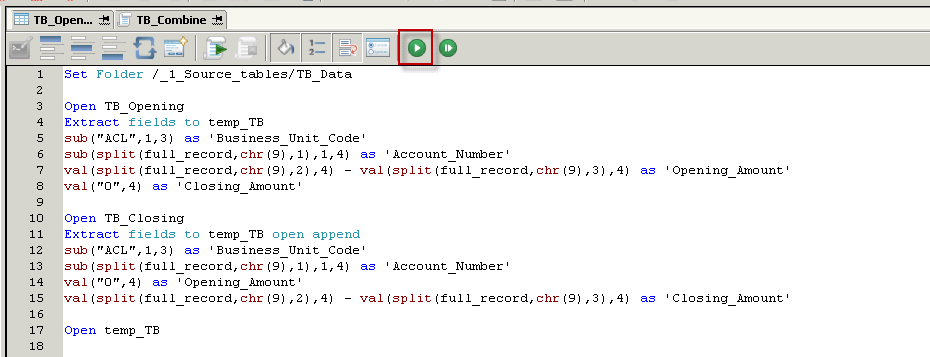
VAL("0", 4) as 'Opening\_Amount'

VAL(SPLIT(full\_record, chr(9), 2), 4) - VAL(SPLIT(full\_record, chr(9), 3), 4) as 'Closing\_Amount'

As shown in Figure 6: Running Script and Temp TB Table, which also shows the full code, you can then run the script by clicking the play button.

Appendix (continued)

Figure 6: Running Script and Temp TB Table



1. Creating the Final TB Table

We have now created the temp\_TB table. However, as is pointed out in Figure 6: Running Script and Temp TB Table, two rows come from the headings in the source data. If we were to scroll down in this table, we would see the same thing again where the closing TB begins. We will return to this situation in a moment.

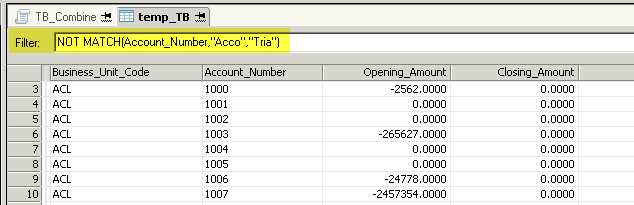
Summarize on Business\_Unit\_Code Account\_Number subtotal Opening\_Amount

Closing\_Amount to TB\_Final open presort

Appendix (continued)

We are now using the summarize command to bring the fields from temp\_TB into a new table, which we call TB\_Final. For this table, we want to remove the title rows of data that we just saw in the temp\_TB table. To do this, we filter using the NOT MATCH(comparison\_value, test1, test2, <test3...>) function, shown in Figure 7: Not Match Function and Filter below. In this function, the comparison\_value argument refers to the column that we want to filter on, and the test arguments refer to the cells that we want to filter out. This function can also be used as simply MATCH(comparison\_value, test1, test2, <test3...>) for the opposite result.

Figure 7: Not Match Function and Filter



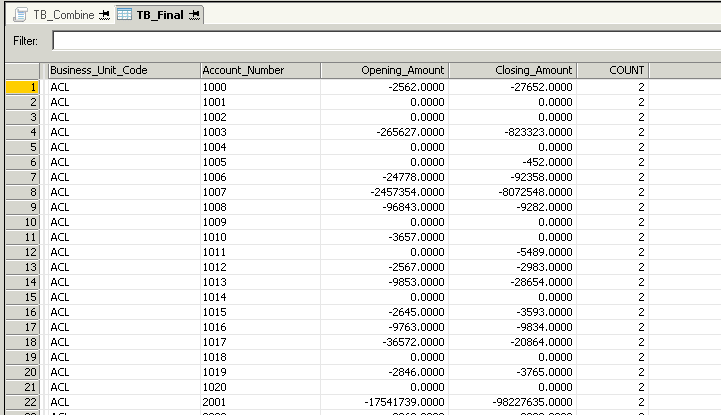
Summarize on Business\_Unit\_Code Account\_Number subtotal Opening\_Amount

Closing\_Amount to TB\_Final open presort if NOT MATCH(Account\_Number, "Acco", "Tria")

We now have a complete script to combine the two TB tables. (The entire script is below under E. Complete Script. Running this script will give us the TB\_Final table (see Figure 8: Final TB Table), which is the product that can be used to reconcile against journal entries in the GL for ICW.

Appendix (continued)

Figure 8: Final TB Table



1. Complete Script

Set Folder /\_1\_Source\_tables/TB\_Data

Open TB\_Opening

Extract fields to temp\_TB

SUB("ACL", 1, 3) as 'Business\_Unit\_Code'

SUB(SPLIT(full\_record, chr(9), 1), 1, 4) as 'Account\_Number'

VAL(SPLIT(full\_record, chr(9), 2), 4) - VAL(SPLIT(full\_record, chr(9), 3), 4) as 'Opening\_Amount'

VAL("0", 4) as 'Closing\_Amount'

Open TB\_Closing

Extract fields to temp\_TB open append

SUB("ACL", 1, 3) as 'Business\_Unit\_Code'

SUB(SPLIT(full\_record, chr(9), 1), 1, 4) as 'Account\_Number'

VAL("0", 4) as 'Opening\_Amount'

VAL(SPLIT(full\_record, chr(9), 2), 4) - VAL(SPLIT(full\_record, chr(9), 3), 4) as 'Closing\_Amount'

Open temp\_TB

Summarize on Business\_Unit\_Code Account\_Number subtotal Opening\_Amount Closing\_Amount to TB\_Final open presort if NOT MATCH(Account\_Number, "Acco", "Tria")

Appendix (Continued)

List of Functions and Commands Used in the Tutorial and Other Useful Functions and Commands

String Functions

* SPLIT(string, separator, segment, <text\_qualifier>)
* SUBSTRING(string, start, length) *or*   
  SUB(string, start, length)
* UPPER(string) *and* LOWER(string)

This function simply changes all letters in the string to either uppercase or lowercase, respectively. It is useful for creating visually appealing and consistent output, as the case does not actually affect the utility of a piece of code in ACL.

* EXCLUDE(string, characters\_to\_exclude)

This function excludes the character specified from a string. For example, an account number string may have hyphens separating different segments, and this function could be used to remove those hyphens.

* INCLUDE(string, characters\_to\_include)

This function works the opposite of the EXCLUDE() function. It returns only the characters that are in both the original string and the characters\_to\_include argument.

* REPLACE(string, old\_text, new\_text)

This function simply replaces part of a string with something new. An example use could be replacing the abbreviation “st” with “street” in an address.

* LENGTH(string) *or* LEN(string)

This function reports the length of the string in question (see Figure 9: Example of the LEN Function). This function can be used prior to using the SUB() function, to pick up on any extra spaces that may have appeared in the data.

* ALLTRIM(string) *or* ALL(string)

This function works to remove any extra spaces that have appeared in data fields. However, unlike the SUB() function, this function can be used where different cells in the one column have varying lengths (for example, the account description field).

Appendix (Continued)

Conversion Functions

* VALUE(string, decimals) *or*   
  VAL(string, decimals)
* STRING(number, length, <format>)

Essentially the opposite of the VALUE() function, this function converts numeric data into character data.

* DEC(number, decimals)

This function instructs ACL how many decimals to return for a numerical field.

* ZONED(number, length) *or* ZON(number, length)

This function is useful when formatting date fields. If dates with single digit days or months are in the source data as single digits (i.e., no 0 before the number), this function can be used to make sure all cells are the same length. For example, if a date is written as 3/5/2017, the ZON() function could be used (in conjunction with other functions) to format the cell as 03/05/2017.

Logical Functions

* MATCH(comparison\_value, test1, test2, <test3...>) *or*   
  NOT MATCH(comparison\_value, test1, test2, <test3...>)

Date and Time Functions

* CTOD(string/number, <format>) *and* CTODT(string/number, <format>)

These functions are used to convert character or numerical data into date or datetime format. CTOD is an abbreviation for “character to date” and CTODT is an abbreviation for “character to datetime.” The <format> argument must be used for any string or value in which the data is in any format other than YYYYMMDD or YYMMDD.

* DAY(date/datetime) *and* MONTH(date/datetime)

These functions are used to separate the specified value from a field that is in date or datetime format. For example, the DAY() function will return the day (from 1 to 31) in numerical format, from the relevant date or datetime data. Other functions of this type include YEAR(), HOUR(), MINUTE(), and SECOND().

Appendix (Continued)

Math Functions

* ABS(number)

This function returns the absolute value of numerical fields.

* ROUND(number)

This function rounds the specified number to the nearest whole number value.

* MAXIMUM(value\_1, value\_2, <value\_3…>) *and* MINIMUM(value\_1, value\_2, <value\_3…>)

These functions return the maximum and minimum values, respectively, of a series of numerical data. They will also return the most recent (MAXIMUM()) or the oldest (MINIMUM()) when used with datetime data.

Commands

* Open *and* Open Append
* Extract Fields
* Summarize *or* Sum
* Comment *or* Com

This command instructs ACL that the following section of text is not code; rather, it is a comment or a description of the code. This command is extremely useful in creating clear code that can easily be used by other people.

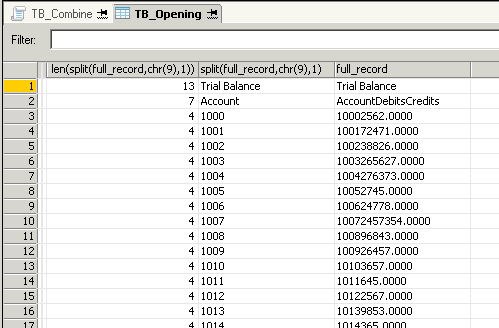
* Statistics

This command calculates various useful statistics on numeric or datetime fields, such as minimum and maximum values, averages, and quartiles.

* Total

This command simply calculates the total value of the specified numeric fields.

Figure 9: Example of the LEN Function



Source: Created by authors.

1. Jonathan Shaw, “Why ‘Big Data’ Is a Big Deal,” *Harvard Magazine*, March–April 2014, accessed September 2, 2017, https://harvardmagazine.com/2014/03/why-big-data-is-a-big-deal. [↑](#footnote-ref-2)
2. Steve LaValle, Eric Lesser, Rebecca Shockley, Michael Hopkins, and Nina Kruschwitz, “Big Data, Analytics and the Path from Insights to Value,” *MIT Sloan Management Review* 52, no. 2 (2011): 21–31. [↑](#footnote-ref-3)
3. Adrian Gepp, Martina Linnenluecke, Terry O’Neill and Tom Smith, “Big Data Techniques in Auditing Research and Practice: Current Trends and Future Opportunities,” *Journal of Accounting Literature* 40 (2018): 102–115. [↑](#footnote-ref-4)
4. The Big Four accounting firms are Deloitte, Ernst & Young (EY), KPMG, and Pricewaterhouse Coopers (PwC). See, for example, Deloitte, “Deloitte Analytics,” accessed July 1, 2017, https://www2.deloitte.com/au/en/pages/technology/solutions/ deloitte-analytics.html; EY, “Becoming an Analytics-Driven Organisation,” accessed July 1, 2017, www.ey.com/gl/en/issues/business-environment/ey-analytics; KPMG, “Analytics, Information and Modelling,” accessed July 1, 2017, https://home.kpmg.com/au/en/home/services/advisory/management-consulting/data-analytics-information-modelling.html; PwC, “Data and Analytics,” accessed July 1, 2017, https://www.pwc.com/gx/en/issues/data-and-analytics.html. [↑](#footnote-ref-5)
5. Kevin Moffitt and Miklos Vasarhelyi, “AIS in an Age of Big Data,” *Journal of Information Systems* 27, no. 2 (2013): 1–19. [↑](#footnote-ref-6)
6. Nadia Carlin, “Getting a Grip on Data Assurance,” PwC (blog), December 2014, accessed July 1, 2017, www.pwc.com.au/publications/audit-committee-matters/data-assurance-dec14.html. [↑](#footnote-ref-7)
7. IAASB Data Analytics Working Group, “Exploring the Growing Use of Technology in the Audit, with a Focus on Data Analytics,” (submission to the Board of Directors, International Auditing and Assurance Standards Board [IAASB], August 25, 2016), accessed November 25, 2016, https://www.ifac.org/publications-resources/exploring-growing-use-technology-audit-focus-data-analytics. [↑](#footnote-ref-8)
8. Ibid. [↑](#footnote-ref-9)
9. Ibid. [↑](#footnote-ref-10)
10. Carlin, op cit. [↑](#footnote-ref-11)
11. Greg Richins, Andrea Stapleton, Theophanis Stratopoulos, and Christopher Wong, “Big Data Analytics: Opportunity or Threat for the Accounting Profession?,” *Journal of Information Systems* 31, no. 3 (2017): 63–79. [↑](#footnote-ref-12)
12. Bernard Marr, “Big Data, AI and the Uncertain Future for Accountants,” *Forbes*, October 7, 2016, accessed August 1, 2017, https://www.forbes.com/sites/bernardmarr/2016/10/07/big-data-ai-and-the-uncertain-future-for-accountants. [↑](#footnote-ref-13)
13. Richins et al., op. cit. [↑](#footnote-ref-14)
14. Julita, “Difference between ACL and IDEA,” DifferenceBetween.net, June 2, 2010, accessed March 9, 2018, www.differencebetween.net/technology/difference-between-acl-and-idea. [↑](#footnote-ref-15)
15. Carlin, op. cit. [↑](#footnote-ref-16)
16. Robert Braun and Harold Davis, “Computer-Assisted Audit Tools and Techniques: Analysis and Perspectives,” *Managerial Auditing Journal* 18, no. 9 (2003): 725–731. [↑](#footnote-ref-17)
17. Ibid. [↑](#footnote-ref-18)
18. Richins et al., op. cit. [↑](#footnote-ref-19)
19. Richins et al., op. cit; Gepp et al., op. cit; Miklos Vasarhelyi, Alexander Kogan, and Brad Tuttle, “Big Data in Accounting: An Overview,” *Accounting Horizons* 29, no. 2 (2015): 381–396. [↑](#footnote-ref-20)
20. IAASB Data Analytics Working Group, op. cit. [↑](#footnote-ref-21)