

**Integrating Data Analytics into the Financial Statement
Auditing Class in an Efficient and Effective Manner Using IDEA**

Allen Blay
Florida State University

Tim Louwers
James Madison University

David Sinason
Northern Illinois University

Jerry Strawser
Texas A&M University

Jay C. Thibodeau
Bentley University

January, 2016

**DISCUSSION DRAFT
COMMENTS WELCOME**

Integrating Data Analytics into the Financial Statement Auditing Class in an Efficient and Effective Manner Using IDEA

It is absolutely essential that accounting professors recognize that the learning goals and objectives for the financial statement auditing course(s) are not static. One major trend in auditing practice is the use of data analysis tools to take advantage of the growing amounts of data available to auditors. This data can take many forms, from client financial data, to images and words from public sources.¹ In a recent (2015) monograph, PricewaterhouseCoopers (PwC) lists five 'new' skills that will be required of auditors moving forward. Although many of these skills require special statistical or programming knowledge, the first listed skill is one that is applicable to all auditors, "Research and identify anomalies and risk factors in underlying data." Although extraction and analysis from client accounting data are critical skills for newly licensed CPAs, we are unaware of sufficient materials to assist professors in integrating data analytics into the auditing classroom. Thus, the goal of this paper is to provide a clear and implementable method to integrate a leading data analysis tool, the IDEA data analysis software, into the auditing class.

Big Data

The auditing environment is rapidly being transformed into an environment characterized by the availability of "Big Data." Among other changes, there has clearly been an increase in the use of "smart" information technology applications which have helped to fuel the increased use of data and analytical tools by entry level audit professionals on the audit. As a direct result, entry level audit professionals are being asked to join public accounting firms having completed coursework that has exposed them to the use of big data and analytical tools on the audit.

Critical issues for our students to learn about include how to identify the right data to analyze and how to present data analyses in a compelling format when documenting work. These steps are important to learn before implementing analyses of data using cutting-edge analytical tools. While analytical tools can rely on data sources that are both internal and external to the client, our understanding is that staff audit professionals in today's environment need to first learn how to make the best use of *internal* client data to provide useful insights to the audit team.

Our Approach to This White Paper

We collectively believe that auditing professors have a unique opportunity to help our students take advantage of opportunities that have emerged as a result of the important trends that have impacted the current public accounting environment. If firms that are hiring our students are working hard to emphasize the importance of

¹ In 2013, it was estimated that up to 90% of all available data in the world was created in the previous two years. (SINTEF2013.)

big data and documentation of skeptical professional judgments, we set out to create a resource paper that was specifically designed to help address these areas in the financial statement audit classroom in an efficient manner.

The purpose of this white paper is to provide a primer on how to effectively integrate “big data” in the financial statement audit classroom. In addition, the white paper is designed to help show students how to effectively document their conclusions in a “big data” environment. In so doing, we focus on the use of a leading data analysis tool (IDEA Data Analysis Software). We collectively believe that IDEA provides an outstanding platform to illustrate the initial steps that need to be taken by our students related to data and analytics while completing the financial statement audit.

To guide our integration of IDEA, we consulted a recent (2015) monograph published by PwC entitled, “Data Driven: What Students Need to Succeed in a Rapidly Changing Business World,” which clearly indicated that the skills needed by entry level auditing and tax professionals will have to include a proficiency in data analytics and technology. For example, insights from a CEO survey discussed in the monograph reveal that

Businesses are preparing for a future that’s different from today. And they expect their talent to adapt. One implication of this rapidly changing business environment is clear—today’s accounting curriculum should be updated to equip students with new skills, especially in technology and data analytics. (p. 4)

More specifically, as it relates to the audit process, PwC (2015) argues that

Data analytics are altering the way the audit process is done at both the transaction and general-ledger levels. Auditors have new tools to extract and visualize data, allowing them to dig into larger, non-traditional data sets and perform more intricate analysis. These data sources aren’t necessarily numbers. Non-traditional data, like images or words, can be found in different forms and pulled from varied sources, so the ability to analyze all of it leads to better insight. Accounting professionals can also use data visualization tools to help others better understand what the data is telling them, such as depicting the ebb and flow of online conversations around a particular topic, or using an interactive chart to allow a user to change inputs and see a new view automatically. (p. 6)

Overall, it seems clear that the world has changed and our accounting curriculum will also need to change to adapt to the new world that our students are about to enter. Based on the PwC monograph, we identified one specific “new” skill listed by PwC as a learning objective that could most clearly be implemented into a core auditing curriculum as a first step to integrating data analytics into the auditing

classroom. Specifically students should be able to “research and identify anomalies and risk factors in underlying data.”² Importantly, while the PwC report is published by only one firm, our conversations with professionals and observations indicate the other large firms (e.g., KPMG, Deloitte, and EY) are largely in agreement with the learning objectives in the document.

Having identified a key learning objective from the PwC monograph, we next set our minds on how to best develop materials to best integrate big data and analytics into the auditing curriculum. To do so, we sought to understand what would be considered “state-of-the-art” data analytics techniques that are currently in high demand.

What we found is that the use of IDEA is one of the best ways for an entry level auditing professional to begin their journey into the world of big data. Simply stated, in today’s auditing environment, big data is manifested in the financial statement auditing process through the use of tools like IDEA. As a result, we believe that this will be an efficient way to move forward in an effective manner. Indeed, IDEA is a tool that is used by many audit firms, and all firms use tools similar to IDEA. Further, IDEA is readily available to students through an academic partnership, and IDEA provides a workbook and data sets as a simple starting point for introducing students to key concepts in data analytics in auditing. There are of course other options for implementing data analytics into an auditing classroom, including competing products (ACL) and visualization tools (Tableau), among many others. However, we chose IDEA as a first step primarily because of the strong academic partnership, data availability, common usage in practice, and relative ease of implementing it into the classroom.

In writing this white paper, we have attempted to create a manual that can be used by professors as a simple guide of how and where to bring data analytics into the classroom in a meaningful way. Students benefit from the demonstration and “hands on” use of an actual analysis tool used in practice. In addition, we believe that this will be a way for auditing professors to show their colleagues and students that they are responding to the “big data” challenge in the audit classroom.

Exhibit 1 shows how we have integrated the current IDEA data set and IDEA Workbook problems (which are quite extensive) as a tool for you to both effectively and efficiently bring this tool into your classroom. We are hopeful that professors will consider using Exhibit 1 as a “road map” that is designed to help identify specific auditing areas where they can integrate IDEA in the classroom. Given the limited time in a typical auditing class, we are not suggesting that professors implement all the potential exercises. Rather, we are attempting to provide instructors with a guide and listing of the available options, and some information about the time the various assignments would take to demonstrate or assign, as well as some extensions beyond the IDEA workbook. Throughout this paper, we have

² PWC (2015), page 11

linked topics to a traditional audit area, a specific IDEA assignment, as well as to specific pages and exhibits in the Louwers et al. 6th edition textbook (2015) to demonstrate where we introduce the various assignments in our classrooms.

Exhibit 1 – Guide to Integrating IDEA into the Auditing Classroom

Topic	IDEA Assignment	Louwers Textbook Reference
Inherent Risk Assessment – Accounts Receivable	Assignment 1 – Compute aging of A/R	Ch. 3 – pages 98-100; Ch. 4 – pages 141-149; Ch. 7
Inherent Risk Assessment - Inventory	Assignment 2 – Identify potentially obsolete inventory	Ch. 3 – pages 98-100; Ch. 4 – pages 141-149; Ch. 9 – pages 385-394
Inherent Risk Assessment - Inventory	Assignment 3- Calculate profit margins	Ch. 3 – pages 98-100; Ch. 4 – pages 141-149
Internal Control Evaluation	Assignment 4 – Exception testing	Ch. 5 – pages 194-198; Ch. 7 – pages 281-284
Internal Control Evaluation - Purchasing	Assignment 5 – Identify payments to unauthorized suppliers	Ch. 5 – pages 194-198; Ch. 8 – pages 327-331
Revenue and Collection Cycle	Assignment 6 – Import client data and reconcile to G/L	Ch. 7 – pages 285-293
Revenue and Collection Cycle – Control Evaluation	Assignment 7 – Authorization control exceptions	Ch. 7 – pages 281-284, Exhibit 7.5
Acquisition and Expenditure Cycle	Assignment 8 – Import client data and reconcile to G/L	Ch. 8 – pages 331-337
Acquisition and Expenditure Cycle – Control Evaluation	Assignment 9 – Testing numerical sequence	Ch. 8 – pages 327-331, Exhibit 8.4
Acquisition and Expenditure Cycle – Detail Testing	Assignment 10 – Testing for unusual and duplicate payments	Ch. 8 – pages 331-337
Production Cycle	Assignment 11 – Import client data and reconcile to G/L	Ch. 9 – pages 385-394
Production Cycle – Valuation Assertion	Assignment 12 – Identify unusual and obsolete inventory items (partial overlap with Assignment 2)	Ch. 9 – pages 385-394
Forensic Evaluation	Assignment 13 – Applying Benford’s Law	Ch. 8 – pages 339-342; Mod. D – pages 701-705
Attribute Sampling	Example 1 - Aircon	Mod. F, pages 754-770
Monetary Unit Sampling	Example 2 – Rice Company	Mod. G, pages 798-811
Monetary Unit Sampling	Assignment 14 – Comprehensive sampling exercise	Mod. G, pages 798-811

The remainder of this white paper is organized around the potential application of IDEA and data analysis techniques in the following topical areas based on Exhibit 1. Throughout the remainder of this white paper, we will make numerous references to both the IDEA Workbook Version Nine (IDEA Workbook) and Louwers et al. *Auditing & Assurance Services*, 6th Edition (Louwers):

Engagement Planning (Chapter 3, Louwers) and Management Fraud and Audit Risk (Chapter 4, Louwers)

- Students will learn about the potential usefulness of improved data analytic techniques during preliminary analytical review procedures (Louwers, pages 141-145). The discussion features a visualization of a text analysis examining the words most commonly used in social media to describe an audit client's newest products during the year under audit, which might be used by auditors to help develop an expectation for revenue.
- Students will learn how IDEA can be used to leverage internal client data to allow for more robust trend analysis (i.e., year over year). Examples include an aging of accounts receivable analysis, a slow-moving inventory analysis, and an analysis of profit margins.

Internal Control Evaluation (Chapter 5, Louwers)

- Students will learn about how advances in information technology have led to an increase in the number of control tests that are now able to be effectively applied to the entire population of control occurrences, as compared to just a sample of control occurrences.
- Students will learn how IDEA can be used to perform exception tests in the accounts receivable and accounts payable cycle.

Sales and Accounts Receivable Cycle (Chapter 7, Louwers)

- Students will learn how to use IDEA to test the existence, completeness, valuation, and accuracy of accounts receivable.
- Students will learn how to test if an account is mathematically accurate by comparing values to summary totals.
- Students will learn how to perform exception tests with an example of identifying large or old items, and balances that exceed credit limits. The use of IDEA to perform extractions through an equation editor is demonstrated, as well as the task of combining multiple sets of data for joint analysis.
- Students will learn how to use IDEA for performing aging analysis.
- Students will test their skills by performing other relevant exception tests not demonstrated in the examples.

Expenses and Accounts Payable Cycle (Chapter 8, Louwers)

- Students will learn the importance of exception testing in audits of accounts payable, and the use of IDEA for performing tests to detect irregular

payments, double-processed invoices, missing checks, and purchases from unauthorized suppliers.

- Students will test their fraud investigation and IDEA skills through an exercise created to detect whether early and late payments have any systematic patterns.

Inventory Cycle (Chapter 9, Louwers)

- Students will learn the importance of analysis that uses dates to evaluate unusual, obsolete, excessive and slow moving inventory, and the use of IDEA for performing tests to extract data and to manipulate data to determine excessive inventory.
- Students will test their skills through an exercise designed to demonstrate how unusual items may be identified for further testing, including possible fraud investigation.

Forensics (Module D, Louwers)

- Students will learn how to apply Benford's Law to identify unusual items in accounts payable and determine whether further investigation is required.
- Students will test their skills through an exercise requiring testing of unusual items in accounts receivable to determine if additional investigation is required.

Audit Sampling (Modules F and G, Louwers)

- Students will learn about the use of IDEA in determining the sample size and evaluating sample results for the tests of controls for AirCon Company shown in Module F (Louwers).
- Students will learn about the use of IDEA in determining the sample size and evaluating the sample results using Monetary Unit Sampling (MUS) for the examination of Rice Company's accounts receivable shown in Module G (Louwers).
- Students will test their skills through a comprehensive MUS example using a "Sample-Detailed Sales" file that requires students to evaluate the nature of the population, determine the sample size, select sample items, and evaluate sample results.

In order to complete the assignments in this white paper, you will need to have access to the IDEA software and the electronic version of the IDEA Workbook. You will also need to have access to the data files for the audit procedures to be completed for Accounts Receivable, Accounts Payable, and Inventory. For each set of files, you will need to import the data into the IDEA software. To take this action, please refer to the following pages in the IDEA Workbook:

- Accounts Receivable: Please follow the steps on pages 26-41 to properly import the sample data files for accounts receivable into IDEA.

- Accounts Payable: Please follow the steps on pages 97-121 to properly import the sample data files for accounts payable into IDEA.
- Inventory: Please follow the steps on pages 189-202 to properly import the sample data files for inventory into IDEA.

ENGAGEMENT PLANNING (CHAPTER 3, LOUWERS) AND MANAGEMENT FRAUD AND AUDIT RISK (CHAPTER 4, LOUWERS)

Preliminary Analytical Procedures

Professional standards require auditors to perform analytical procedures during the planning stages of the audit “with the objective of identifying unusual or unexpected relationships” involving significant financial accounts “that might indicate a material misstatement, including material misstatement due to fraud.” When doing so, the auditor should consider all types of relevant data to help improve their understanding of risk on the audit. Importantly, professional standards allow the use of “data that is preliminary or data that is aggregated at a high level” when completing analytical procedures at the planning stages. As a result, the increased use of data and analytical tools has the potential of improving the effectiveness of this type of risk assessment procedures.³

Big Data

Indeed, the auditing environment is rapidly being transformed into an environment characterized by the availability of “Big Data.” For example, as it relates to the audit process, PwC’s (2015) document on what students should know about data argues that “Data analytics are altering the way the audit process is done . . . auditors have new tools to extract and visualize data, allowing them to dig into larger, non-traditional data sets and perform more intricate analysis.” There is perhaps no better way to illustrate the potential usefulness of improved data analytic techniques than its use while completing analytical review procedures during the planning stages of the audit. As stated in the PwC (2015) monograph,

These data sources aren’t necessarily numbers. Non-traditional data, like images or words, can be found in different forms and pulled from varied sources, so the ability to analyze all of it leads to better insight. Accounting professionals can also use data visualization tools to help others better understand what the data is telling them, such as depicting the ebb and flow of online conversations around a particular topic, or using an interactive chart to allow a user to change inputs and see a new view automatically.⁴

³ See PCAOB Auditing Standard No. 12 – Identifying and Assessing Risks of Material Misstatement, paragraphs 46-48.

⁴ PwC (2015), page 6

External and Internal Data

An important example of the use of additional data involves the use by auditors of more predictive analytical tools that utilize third party data to supplement their “traditional” analytical procedures. The additional data can help auditors refine their expectations and improve the results of preliminary analytical procedures which form initial beliefs about the nature, timing and extent of audit evidence to be gathered at an audit client. While this type of access to increased volumes of data on the client has the potential to improve audit effectiveness, it also can have an initial negative impact on audit efficiency if audit professionals are unable to efficiently execute such additional procedures.

Overall, it seems clear that the auditing world has changed and when students enter the profession, expertise in the use of data analytics during analytical procedures will be helpful. Interestingly, while the additional analytical tools described above primarily involve data sources that are external to the client, audit professionals in today’s environment also need to learn how to make the best use of internal client data when completing analytical procedures completed during the planning stages.

For example, when completing preliminary analytical procedures, the availability of largely all of the client’s internal data can allow for more robust trend analysis (i.e., year over year) on a multitude of financial and non-financial data.

IDEA

The IDEA software package may also be helpful when summarizing internal client data for purposes of analytical procedures used during the planning process. For example, the IDEA Workbook notes that

IDEA can help with the preparation of figures for an analytical review. In particular, IDEA can generate analyses that would not otherwise be available. The **Stratification** task (from the **Analysis** tab on the IDEA Ribbon) generates a profile of the population in value bands, groups of codes, or dates. This is particularly useful when auditing assets such as accounts receivable, inventories, loans, or for a breakdown of transactions. Additionally, the information can be summarized by particular codes or sub-codes. Figures can also be compared against previous years to determine trends. A chart can be produced if required. (p. 16)

The following exercises can be completed to illustrate the use of IDEA during preliminary analytical procedures. To be most useful, each of these analyses would have to be completed for multiple years so comparisons could be made and meaningful expectations could be developed.

Assignment 1: Conduct an Aging Analysis of Accounts Receivable (Estimated time = 10-15 min): The production of an aging schedule is a common analysis that is performed to help better understand inherent risk for the valuation assertion for accounts receivable. This exercise could be assigned to illustrate how comparisons might be made for each of the aging categories (e.g., 30-60 days, 61-90 days, over 90 days) over the past several years. Large fluctuations year over year might reveal a situation that should be investigated further by the auditor. This exercise also introduces students to the graphing functions in IDEA. Students should complete the exercise in the IDEA Workbook that can be found on pages 53-57.

Assignment 2: Identify Potentially Obsolete Inventory Items (Estimated time = 10-15 min): The identification of the oldest inventory items in stock is a common analysis that is performed to help better understand inherent risk for the valuation assertion for inventory. This exercise could be assigned to illustrate how comparisons could be made for the total cost of “old” inventory items in stock over the past several years. Large fluctuations year over year might reveal a situation that should be investigated further by the auditor. Students should complete the exercise in the IDEA Workbook on pages 203-206.

Assignment 3: Analyze Profit Margins (Estimated time = 20-25 min): The calculation of profit margins by product is an analysis that is performed to help better understand inherent risk for the valuation assertion for inventory. In addition, the analysis could be used to help identify whether a fraud risk exists for the audit client. This exercise illustrates how comparisons could be made for the profit margin by product over the past several years. Large fluctuations year over year might reveal a situation that should be investigated further by the auditor. Students should complete the exercise in the IDEA Workbook on pages 222-230.

INTERNAL CONTROL EVALUATION (CHAPTER 5, LOUWERS)

Tests of Internal Control

In today's auditing environment, an auditor now has far more data available to be used for such tests. Recall that after gaining an understanding of internal controls, if the auditor intends to assess control risk as low and rely on a control activity, the auditor needs to gather evidence to confirm that the control activity has been designed and has operated effectively during the entire period of control reliance. This is necessary because a primary factor in determining the nature, timing, and extent of substantive procedures to be performed is a lower control risk assessment. Simply stated, substantive procedures are typically reduced as a result of reliance on well-designed internal controls that are operating effectively. In today's auditing environment, it is critical that audit professionals learn how to make the best use of internal client data when designing and completing the tests of controls needed to support the conclusion that a control activity has been designed and has operated effectively during the period of reliance.

The professional standards make clear that when designing tests of controls, the auditor needs to consider the means of selecting items for testing. For tests of internal controls, there are two approaches that are commonly used: (1) testing all items in a population; and (2) testing a sample from a population. The decision of which approach to use depends on the nature of the control that is being tested, along with the availability of data. For example, a control activity that is entirely automated might best be tested by an automated audit procedure that can be efficiently and effectively be applied to the entire population of occurrences of that control activity. However, for a manual control activity, the auditor is likely to take a sample from the population of occurrences of that control activity. In addition, it should be noted that some manual controls (such as locking a door to safeguard assets) may have little documentation and may require even other means of testing (e.g. observation and inquiry). In today's auditing environment, there has been a dramatic increase in the number of control tests that are able to be effectively applied to the *entire* population of control occurrences in an efficient manner.

Exception Testing

One way to subject all items in a population of occurrences for a particular control activity is to use exception testing. Exception testing is designed to identify a violation of a particular control activity through the use of an automated test procedure designed to test all items in a population. For example, consider an entirely automated control activity that is designed to compare a customer's credit limit to the sum of (1) a potential sales transaction; and (2) that customer's outstanding credit balance before approval of that sales transaction. If the control activity operated effectively throughout the year, a customer's outstanding credit balance would not exceed its credit limit.

Given the nature of the control activity, one way to test the operating effectiveness would be through the use of exception testing. That is, an auditor could obtain evidence about the control's operating effectiveness by using a procedure that compares each customer's credit limit to that customer's outstanding credit balance at the end of each day for the year under audit. Such a testing strategy would not have been possible (at least economically) in previous years. However, due to advances in information technology, such testing is now possible. As a direct result, entry level audit professionals are now expected to consider the full extent of client data available for testing purposes, before proceeding with audit tests.

Audit Sampling

Of course, there are many control activities that do not lend themselves to automated audit testing. In such situations, auditors are likely to take a sample from the population of occurrences for the control activity being tested. Most importantly, in such situations, the population being sampled must include all occurrences of the relevant control activity for the entire period of reliance and the

sample must be representative of that population to be considered appropriate audit evidence. In recent years, there has been an increased focus on the work of auditors by regulators to ensure that the sample selected is truly representative of the population of occurrences over the entire period of reliance. For example, in a recent Staff Practice Audit Alert, the PCOAB (2014) noted that its inspectors observed instances where auditors relied on controls to reduce substantive testing but their reliance was “unsupported because the testing of controls was insufficient” because auditors failed to test the control activity over the entire period of reliance.⁵ In such situations, the key is for auditors to remember that for a sample to be representative; all items in the population have an opportunity to be selected.

IDEA

IDEA can be helpful to audit professionals when completing exception tests and conducting audit sampling. The following assignments allow students to experience a hands-on application of the IDEA software to exception testing. Ultimately, exception tests provide evidence about the operating effectiveness of internal control activities by testing all items in a population. Later in this white paper, we illustrate how auditors use the sampling features of IDEA to select a representative sample from a complete population of control occurrences for a control activity to be tested.

Exception Test Assignments

Assignment 4: Extracting Records to Identify Authorization Control Exceptions:

The proper authorization of customer credit is an example of a control activity where the operating effectiveness can be tested using an exception test run with IDEA. We recommend assigning this exercise in conjunction with a classroom discussion of the importance of the control activity that is designed to ensure the proper authorization of customer credit limits. In order to complete the assignment, the following two steps need to be completed:

Step 1: Analyze the Balances by Customer Account (Estimated time = 15-30 min): To begin this assignment, the balances by individual customer must be summarized and a separate report that can be linked to the customer credit authorization file must be created. Following the instructions on pages 68-70 of the IDEA Workbook, create the customer balance summary. The instructions in the IDEA Workbook are clear, but it should be noted that the column NET referred to in the instructions does not exist since it was created in an exercise that has not been completed, and it is not necessary to complete this exercise.

⁵ From “Staff Audit Practice Alert No.12 Matters Related to Auditing Revenue in an Audit of Financial Statements,” PCAOB, September 9, 2014. Available at http://pcaobus.org/Standards/QandA/9-9-14_SAPA_12.pdf.

Step 2: Evaluate Customer Balances against Authorized Credit Limits (Estimated time = 30-60 min): This exercise is more complex than previous examples and requires a client file to be imported and merged with another file with different data fields. However, these requirements are quite common whenever combining client data that may not have similar formats or field names for the matching columns (keys).

Follow the instructions on pages 71-90 of the IDEA Workbook. In many instances in this exercise, students will need to “hover” over icons in the menu to find the correct functions (that are listed by name in the IDEA Workbook). In addition, step 7 on page 76 of the IDEA Workbook instructs the user to input a Line Length of five characters. While it is not possible to specify this, if all other items are input correctly, the address field is correctly highlighted and the field editor should be working correctly.

The assignment provides significant practice performing direct data extractions, and the IDEA Workbook does not always provide direct guidance. Remember this task because it will be used in future assignments.

Assignment 5: Identifying Payments to Unauthorized Suppliers

The proper authorization of suppliers (or vendors) is an example of a control activity where the operating effectiveness can be tested using an exception test run with IDEA. We recommend performing this exercise in conjunction with a classroom discussion of the importance of the control activity that ensures that payments are only made to authorized suppliers (or vendors). In order to complete the assignment, the following two steps need to be completed:

Step 1 (Estimated time = 15-30 minutes): To complete this step, the disbursements database must be merged with the authorized supplier database, as described in pages 171 – 177 (Step 19) of the IDEA Workbook. Although the concepts are complex, including elimination of duplicate keys, the explanations in the IDEA Workbook are reasonably straightforward and easy to follow.

Step 2 (Estimated time = 5-10 minutes): Once the databases have been merged in Step 1, identifying payments that are not matched to a supplier is relatively straightforward. By following Steps 20-23 on pages 177-178 of the IDEA Workbook, one supplier (M100) is discovered to have multiple payments.

REVENUE AND COLLECTION CYCLE (CHAPTER 7, LOUWERS)

For most students, the revenue cycle is the first full cycle covered in an auditing course. For that reason, the audit of the revenue cycle, in particular sales and accounts receivable, is an excellent opportunity to introduce students to handling a

data set in IDEA, and addressing some common issues that auditors face when importing client data. This addendum summarizes several useful exercises available to students in the IDEA Workbook using the IDEA datasets, and provides instructors with an additional exercise to assign students to test their understanding and proficiency with analyzing client data in IDEA. We suggest that if instructors only wish to introduce students to IDEA in one cycle, following the exercises in this section will sufficiently introduce students to many of the data extraction, importation, and manipulation tools available in IDEA that can be applied to other cycles as well.

Assignment 6: Importing Client Data and Reconciling to General Ledger

This assignment must be completed prior to any of the other assignments in this section. Other assignments in this section are designed to be independent of each other, but this assignment should be considered as a prerequisite for the remaining assignments.

Step 1 (Estimated time = 15-30 Minutes): Students should first follow the instructions on pages 26-41 of the IDEA Workbook. The instructions are clear and little instructor guidance should be needed (although the file location of the My IDEA Documents directory may require students to enter the My Documents folder first). The students will discover that the data imported into IDEA (\$435,864.85) does not agree to the totals provided by the client (GROSS_AMT \$369,255.13). This is a common problem encountered in practice, especially in the revenue and expenditure cycles because client supporting files often include data that differs in timing from the financial statements (cut-off problems) and problems with delayed removal of items (existence problems).

Step 2: (Estimated time = 15-30 Minutes): In this step, students will learn the basic direct data extraction tool in IDEA using the Equation Editor. To reconcile the imported data to the client's control totals (G/L), students are instructed that the discrepancy relates to both cut-off (incorrect dates) and existence (some items already marked with a paid flag) issues. By following the instructions on pages 42-49 of the IDEA Workbook, students will be able to successfully reconcile the imported records to the client's records and have a sub-database that will be used for additional assignments.

Note that the "Equation Editor" button referred to in the IDEA Workbook is a calculator image located directly to the right of the File Name field box in the direct extraction box. This is not explained in the instructions and may not be obvious to students. All other instructions should be clear to students with reasonable effort. It is critical that students recognize the importance of learning the formatting requirements for extraction equations in any type of audit software. Indeed, knowledge of the formatting requirements is essential to accurately extracting data.

Assignment 7: Extracting Records to Identify Authorization Control Exceptions: Pages 282-284 of Louwers (particularly Exhibit 7.5) discusses important controls in the revenue cycle. Authorization of credit is one instance where the operating effectiveness can be tested using IDEA. We recommend assigning this exercise in conjunction with a classroom discussion of authorization of credit controls.⁶

Step 1: Analyze the Balances by Customer Account (Estimated time = 15-30 min): To begin this assignment, the balances by individual customer must be summarized and a separate report that can be linked to the customer credit authorization file must be created. Following the instructions on pages 68-70 of the IDEA Workbook, create the customer balance summary. The instructions in the IDEA Workbook are clear, but it should be noted that the column NET referred to in the instructions does not exist since it was created in an exercise that has not been completed, and it is not necessary to complete this exercise.

Step 2: Evaluate Customer Balances against Authorized Credit Limits (Estimated time = 30-60 min): This exercise is more complex than previous examples and requires a client file to be imported and merged with another file with different data fields. However, these requirements are quite common whenever combining client data that may not have similar formats or field names for the matching columns (keys).

Follow the instructions on pages 71-90 of the IDEA Workbook. In many instances in this exercise, students will need to “hover” over icons in the menu to find the correct functions (that are listed by name in the IDEA Workbook). In addition, step 7 on page 76 of the IDEA Workbook instructs the user to input a Line Length of five characters. While it is not possible to specify this, if all other items are input correctly, the address field is correctly highlighted and the field editor should be working correctly.

ACQUISITION AND EXPENDITURE CYCLE (CHAPTER 8, LOUWERS)

Data analytics in the acquisition and expenditure cycle often focuses on exception testing – identifying items that do not seem to fit with expectations. The assignments in this section focus on exception testing, with an example of extracting data using queries that merge multiple data sources. The assignments are ordered to match the flow of the Louwers and do not match the order in the IDEA Workbook. With the exception of Assignment 1, which must be completed to load and verify the data, all assignments are independent and can be selected based on the preferences of the instructor. Instructors may wish to inform students that the database labeled

⁶ This assignment is the same as Assignment 1 in the Internal Control section. Whether an instructor assigns this in coverage of Internal Controls or coverage of the Revenue Cycle is a matter of personal preference.

“Accounts Payable” is actually a listing of disbursements for the period and is not a listing of all invoices or current payables.

Assignment 8: Importing Client Data and Reconciling to General Ledger

Step 1 (Estimated time = 30-45 Minutes): Students should first follow the instructions on pages 100-115 of the IDEA Workbook. The instructions are clear and little instructor guidance should be needed (although file location of the My IDEA Documents directory may require students to enter the My Documents folder first if they have not completed prior exercises). Importing this data involves two data sets, one which is a text document that will require formatting. Students should be cautioned to pay close attention to column formatting in Step 7 on page 107 of the IDEA Workbook, as errors in this step will result in the inability to complete future steps.

Step 2 (Estimated time = 10-15 Minutes): Once files are imported, students will test to ensure that the data is complete and agrees to the client’s control totals. By completing pages 116-121 of the IDEA Workbook, students will conclude that the data agrees to control totals.

Important: The Earliest and Latest Date statistics shown in step 3 on page 118 of the IDEA Workbook do not match to what students will have. The correct dates are 1/4/2012 and 6/12/2012 for the INV_DATE field, and 1/12/2012 and 12/29/2012 for the PAY_DATE field.

Assignment 9: Testing Numerical Check Sequences (Estimated time = 10-15 Minutes)

Pre-numbered documents serve as an important control for the completeness assertion and are specifically discussed in the Louwers (see Exhibit 8.4 on page 329). A critical test when evaluating the operating effectiveness of this control in many areas of the audit is to scan the sequence of documents and identify missing items. IDEA makes this simple and students will learn the **Gap Detection** feature of IDEA by completing pages 154-158 of the IDEA Workbook, which involves searching for missing check numbers. When discussing missing check numbers in class, it is useful to ask students to consider reasons for why checks may be missing and if any patterns exist. In this case, nearly all the missing check numbers are in one block, which may be indicative of incomplete recording of disbursements.

Assignment 10: Searching for Unusual and Duplicate Payments

Step 1 (Estimated time = 15-30 Minutes): Data extraction for exception testing is especially critical in the accounts payable area. In this exercise, students test the payments for high amounts, checks written to “CASH”, payments made on Sunday (when many companies do not conduct business), round dollar amounts, and payments authorized by one individual. IDEA’s

ability to parse text is a rich feature and by completing pages 132-138 of the IDEA Workbook, students have the opportunity to gain experience in formatting data extractions and using text parsing functions. If students have not utilized the Equation Editor, note to them that the **Equation Editor** is the calculator symbol shown in the middle of the data extraction window.

This exercise presents an opportunity to discuss with students what makes a payment “unusual”. The examples of round amounts and Sundays can lead to interesting discussions. An extension to this exercise is to ask students to come up with an example of another unusual type of payment and present the results of their test on this data. One interesting example is to complete an additional extraction that combines facts mentioned in the IDEA Workbook, but not tested – payments authorized by HMV (the initials of an employee) above the authorization limit. If students correctly perform the extraction [`@Upper(@Strip(AUTH)='HMV' .AND. AMOUNT>20000]`, they will obtain control totals of 226 payments totaling \$12,504,294.88. This is an excellent opportunity to discuss how the auditor would proceed in this situation, when it becomes clear that either their understanding of a control is incorrect, or there is potentially a material fraud occurring.

Step 2 (Estimated time = 10-15 Minutes): Another important test in analyzing both controls and the existence of accounts payable is searching for items that are recorded twice (either through double-recording or double-submission of invoices). IDEA has a **Duplicate Key** function that enables auditors to identify duplicate items regardless of the area of the audit. By completing pages 149-153 of the IDEA Workbook, students will have the opportunity to use this feature to search for duplicate invoices. Although the exercise is straightforward, the concept is rich because it demonstrates the importance of evaluating multiple criteria when searching for duplicate payments. The exercise involves both identifying multiple payments of the same amount to the same supplier, and detecting suppliers with checks written to different names. Another interesting feature of the exercise is that it identifies a control weakness in the company that would likely reach the level of at least a significant deficiency and may be indicative of management fraud.

PRODUCTION CYCLE (LOUWERS, CHAPTER 9)

The audit of the production cycle, in particular inventory, is an outstanding opportunity for students to use IDEA to use data analysis to identify items that require further investigation. This addendum summarizes several useful exercises available to students in the IDEA Workbook using the IDEA datasets, and provides instructors with an additional exercise to assign students to test their understanding and proficiency with analyzing client data in IDEA. Assignments 1 and 2 are from the IDEA Workbook and provide step-by-step data analysis.

Assignment 11: Importing Client Data and Reconciling to General Ledger (Estimated time = 15-30 min):

This assignment must be completed prior to attempting other assignments in this section. When importing this file, note that it is in ASCII format, which differs from other IDEA dataset files.

Students should first follow the instructions on pages 193-202 of the IDEA Workbook. The instructions are clear and little instructor guidance should be needed (although file location of the My IDEA Documents directory may require students to enter the My Documents folder first if this has not been done in prior exercises). It is important that the students rename the headings on each column to match the headings in Step 7 on page 196 of the IDEA Workbook. In addition, the column type and number of decimals (for numeric columns) must be entered. For date columns, the date format (YYYYMMDD) must also be entered. The students should verify the data are correct as indicated on pages 200 – 202 of the IDEA Workbook.

Assignment 12: Extracting Records to Identify Unusual and Obsolete Inventory items: The valuation assertion is one of the most critical elements in this cycle and identifying potential unusual and obsolete inventory items is important in reaching an audit conclusion regarding valuation. We recommend assigning this exercise in conjunction with a classroom discussion on identifying obsolete inventory. **Note: If students completed Assignment 2, they will have completed pages 203-206 already.**

Step 1: Analyze Inventory for unusual and obsolete inventory (Estimated time = 20-30 min): Students should follow the instructions on pages 203-213 of the IDEA Workbook to extract information regarding inventory which may be obsolete. Three different extractions are performed in the same process:

1. The client has noted some items as being obsolete. These items are extracted and included in the data analysis.
2. Some items are listed with negative inventory amounts which are clearly an issue and needs further audit investigation. These items are extracted and included in the data analysis.
3. Some items are listed with a total cost that is less than zero. Again, these items clearly need further audit investigation. These items are extracted and included in the data analysis.

The instructor may wish to have a discussion of how the students might use these data and what further audit investigation is required. Note, a discussion of materiality may be useful here. Student may view inventory items with negative amounts or negative costs as immaterial, but the actual amount or actual cost may be significant. Failure to investigate these items would not demonstrate sound professional judgment.

Step 2: Using Data Analysis to Extract Potential Obsolete Inventory Items (Estimated time = 30-60 min): In Step 1 of this assignment, students extracted items that the client flagged as obsolete. It would be a grave error to assume that either the client has identified all obsolete items or that the client has indicated in the data all items that they know to be obsolete. Clearly, obsolete inventory needs to be reduced in value if not written off completely. The client may wish to defer these write downs to a future period in an effort to manage earnings. As such, auditors need to use data analysis to identify potential obsolete items which are unknown to management or items that management may know about but do not wish to communicate to the auditors.

Students should follow the instructions on pages 206-213 of the IDEA Workbook. The first analysis is to calculate the number of months of inventory being carried by the client for each item. Both the *@age* and *@if* functions are used in this analysis. These functions are important and powerful tools in performing meaningful analysis. In addition the *stratification* function is used which can assist auditors in determining materiality levels and in selecting samples that focus on the most significant items.

Note, the final report indicates the months of inventory carried by the client. The exercise indicates that 6 months inventory is acceptable for the client without any additional explanations. This is a good point for discussion regarding how to proceed. Clearly if **Bright IDEAs, Inc.** is, for example, a retail electronic company (e.g., computers, cell phones, tablets, etc.) due to potential obsolescence carrying more than a 1 -2 months inventory may be questionable and a sample of items with inventory on hand greater than 3 months should be investigated. In contrast, a furniture store may have in its warehouses 1 – 6 months' worth of inventory and the auditors would only be concerned with inventory items that represent more than a 6 month supply. Further, if you were a football team, for example the Chicago Bears, having six months' worth of inventory for football jerseys, tee shirts, team logo blankets and other items sold at the stadium would be appropriate at the beginning of the season, but would likely be excess if six months remained at the end of the season. Such a discussion illustrates the importance of industry knowledge in performing an audit, and reinforces the use of data analytics as a tool that does not replace audit judgment.

FRAUD AUDITING (LOUWERS, MODULE D)

One of the tools available to auditors is Benford's Law. Benford's Law identifies suspicious patterns in transactions and events and may indicate when a fraud investigation is warranted. In addition, it is used by fraud examiners to hone in on those items that need evaluation. This addendum summarizes a useful exercise available to students in the IDEA Workbook using the IDEA datasets and allows

students to use the Benford's Law tool to develop information useful in making investigation decisions. In addition, this addendum provides instructors with an additional exercise to assign students to test their understanding and proficiency with Benford's Law in IDEA. The assignment is from the IDEA Workbook and provides a detailed step-by-step data analysis.

Assignment 13: Importing Client Data and Using Benford's Law (Estimated time = 15-30 min):

This assignment uses the accounts payable data set that was used in the assignments in the Acquisition and Expenditure section of this white paper. If the students have completed this assignment, they may skip **Step 1**. However, the importation and preparation of the data set must be completed before the students can complete **Step 2**.

Step 1 (Estimated time = 30-45 min): Students should first follow the instructions that begin on pages 100-115 of the IDEA Workbook. The instructions are clear and little instructor guidance should be needed (although file location of the My IDEA Documents directory may require students to enter the My Documents folder first if they have not completed prior exercises). Students should be cautioned to pay close attention to the column formatting in Step 7 of the IDEA Workbook, as errors in this step will result in the inability to complete future steps.

Step 2: (Estimated Time = 20-40 min): In this step, students will learn how to use Benford's Law in IDEA. By following the instructions on pages 139-149 of the IDEA Workbook, students will successfully collect data on accounts using Benford's Law. This section includes an introduction into the nature of Benford's Law, so little introduction should be required before the students begin this section.

ATTRIBUTE SAMPLING (LOUWERS, MODULE F)

The IDEA Workbook does not contain detailed examples or guidance with respect to the use of the IDEA software in an attribute sampling application. To illustrate the use of IDEA in this regard, consider the AirCon example shown in Module F of Louwers. Because attribute sampling is not included in the IDEA Workbook, this section of the white paper provides detailed guidance and sample screen shots.

Example 1 –Attribute Sampling: AirCon

From the AirCon example, recall the following basic parameters related to the audit team's evaluation of the occurrence control (See Louwers, Exhibit F.6, page 763):

- Risk of overreliance = 5% (which corresponds to a Confidence level of 95%)
- Tolerable rate of deviation = 6%
- Expected population deviation rate = 2%

- Population size = greater than 20,000 transactions

Determining Sample Size (Louwers, pages 760-763)

The Attribute Sampling Module can be accessed through the “Analysis” tab in IDEA. Using the **Analysis>Sample>Attribute** function, the audit team would enter the parameters noted in the “Planning (Beta Risk Control)” tab (Population size, % Tolerable deviation rate, % Expected deviation rate, Confidence level). For population size, enter 20000 (IDEA will not accept commas); while AirCon processes over 20,000 transactions, once a population reaches a certain level, population size has a minimal effect on sample size. After selecting “Compute”, the following output is provided. (NOTE: data input by the audit team are circled in red, IDEA processing results are circled in green)

The screenshot shows the 'Attribute Sampling' window with three tabs: 'Planning (Beta Risk Control)', 'Planning (Beta and Alpha Risk Control)', and 'Sample Evaluation'. The 'Planning (Beta Risk Control)' tab is active. Input fields for 'Population size' (20000), '% Tolerable deviation rate' (6.00), '% Expected deviation rate' (2.00), and 'Confidence level (to control Beta Risk)' (95.00) are circled in red. The 'Sample size' is 127, and the 'Critical number of deviations in sample' is 3. A table shows the relationship between the number of deviations and the achieved confidence level. The row for 3 deviations (2.36% deviation rate) and 95.11% confidence is highlighted in yellow. A green circle highlights the sample size of 127. A conclusion text is provided at the bottom.

Deviations	% Deviations	Achieved Confidence (Beta Risk Control)
0	0.00	99.96
1	0.79	99.66
2	1.57	98.42
3	2.36	95.11
4	3.15	88.53
5	3.94	78.15
6	4.72	64.63
7	5.51	49.69
8	6.30	35.36
9	7.09	23.26

Conclusion: If no more than 3 deviations are observed in a sample of size 127, you can be at least 95.00% confident that the population deviation rate is not more than 6.00%.

This output indicates that:

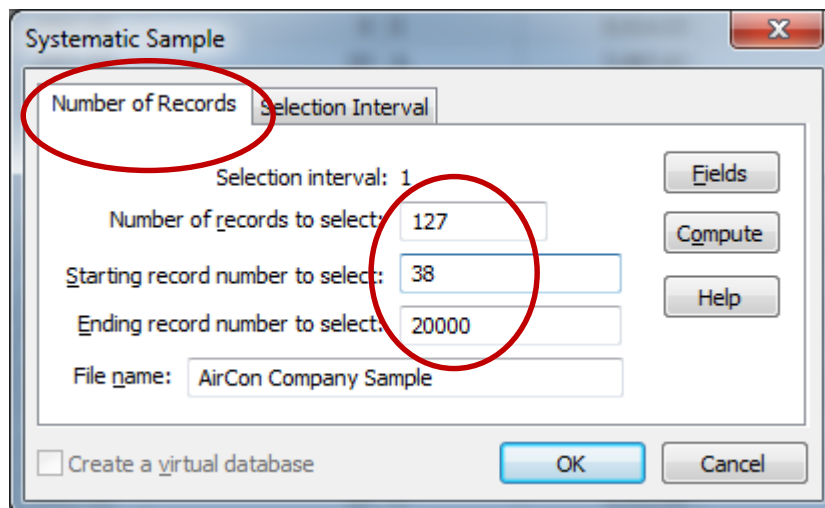
- The appropriate sample size is 127 items (which is consistent with the result using Louwers, Exhibit FA.1, page 791 and summarized on page 761 of Louwers)
- If three or fewer deviations are identified, the audit team will have achieved its desired level of the risk of overreliance of 5%. This corresponds to the number in parentheses shown in the AICPA sampling table reproduced as Exhibit FA.1 (Louwers, page 791).

Selecting the Sample (Louwers, pages 764, 793 - 794)

Once the sample size has been determined, the audit team can use IDEA to select sample items for examination. IDEA can use either unrestricted random selection or systematic random selection techniques with an electronic file of the population, as follows:

- The **Analysis>Sample>Random** function requires the audit team to enter the number of records to select (sample size); the random start, starting record number, and ending record number will all be populated by IDEA. When “OK” is selected, the sample is automatically drawn from the population.
- The **Analysis>Sample>Other>Systematic** function will provide two tabs: “Number of Records” and “Selection Interval”, either of which can be used to select the sample. Depending upon the tab chosen, the audit team will enter the sample size (“Number of records”) or sampling interval (“Selection Interval”) as well as the random starting point in the sample (“Starting record number to select”). IDEA will then select a sample from the population using the random starting point and the sampling interval.

Assuming a population of 20,000 invoices, a desired sample size of 127 items, and a random starting point of 38,⁷ the IDEA input screen would appear as follows (note that this uses the “Number of Records” tab):



The screenshot shows the 'Systematic Sample' dialog box in IDEA. The 'Number of Records' tab is selected and circled in red. The 'Number of records to select' field contains the value 127, the 'Starting record number to select' field contains 38, and the 'Ending record number to select' field contains 20000. The 'File name' field contains 'AirCon Company Sample'. There are buttons for 'Fields', 'Compute', 'Help', 'OK', and 'Cancel'. A checkbox for 'Create a virtual database' is at the bottom left.

By selecting “OK”, IDEA will create a file with 127 records drawn from the population of invoices.

⁷ IDEA does not automatically generate a random starting point; this would need to be determined by the audit team. The **Analysis>Sample>Other>Generate Random Numbers** function can be used to generate a random number for use in an attribute sampling application. However, IDEA does not carry this number forward in the sampling process and it would need to be manually entered by the audit team.

Evaluating Sample Results (Louwers, pages 765 - 767)

Recall that the audit team selected 127 sales invoices and evaluated these invoices for compliance with the control. Also recall that two deviations were identified. At this point, the audit team can evaluate the sample results through the **Analysis>Sample>Attribute** function and selecting the “Sample Evaluation” tab by entering the following parameters:

- Population size = 20000
- Sample size = 127 (from previous step)
- Number of deviations in sample = 2 (based on tests of controls)
- % Desired confidence level = 95% (corresponding to a 5% risk of overreliance)

After selecting “Compute”, the following output is provided (as before, data input by the audit team are circled in red, IDEA processing results are circled in green):

The screenshot shows the 'Attribute Sampling' window with three tabs: 'Planning (Beta Risk Control)', 'Planning (Beta and Alpha Risk Control)', and 'Sample Evaluation'. The 'Sample Evaluation' tab is active. Input fields for 'Population size: 20000', 'Sample size: 127', 'Number of deviations in sample: 2', and '% Desired confidence level: 95.00' are circled in red. Below these, the 'Sample deviation rate: 1.57% (2 / 127)' is displayed. A table shows the results: '1-Sided Upper Limit' is 4.86 (circled in green), '2-Sided Lower Limit' is 0.19, and '2-Sided Upper Limit' is 5.56. A 'Conclusion:' section states: 'For an observed number of 2 deviations in a sample of size 127, you can be 95.00% confident that the population deviation rate is no more than 4.86%. Alternatively, you can be 95.00% confident that the population deviation rate is between 0.19% and 5.56%.' At the bottom, there are buttons for 'Print', 'Close', 'Compute', and 'Help'.

1-Sided Upper Limit	2-Sided Lower Limit	2-Sided Upper Limit
4.86	0.19	5.56

Conclusion: For an observed number of 2 deviations in a sample of size 127, you can be 95.00% confident that the population deviation rate is no more than 4.86%. Alternatively, you can be 95.00% confident that the population deviation rate is between 0.19% and 5.56%.

This output indicates that the upper limit rate of deviation (shown as 1-Sided Upper Limit in the above screen shot) is 4.86% (recall from the AirCon example in Louwers (page 767) that a sample size of 125 was used to correspond to values shown in Exhibit FA.1 and the upper limit rate of deviation was 5.0%). The 1-Sided Upper Limit value is used, as the audit team is conducting a one-tail test; they are concerned that the rate of deviations do not exceed a certain level, not whether the rate of deviations are less than a certain level. Because the upper limit rate of deviation is less than the tolerable rate of deviation, the audit team would conclude that this control was functioning properly.

MONETARY UNIT SAMPLING (LOUWERS, MODULE G)

The IDEA Workbook does not contain detailed examples or guidance with respect to the use of the IDEA software in a variables sampling application.⁸ To illustrate the use of IDEA in a monetary unit sampling application, this white paper illustrates the Rice Company example shown in Module G of Louwers. As variables sampling is not included in the IDEA Workbook, this section of the white paper provides detailed guidance and sample screen shots.

Example 2 - Monetary Unit Sampling: Rice Company Example

From the Rice Company example, recall the following parameters established by the audit team (Louwers, pages 800-802):

- Recorded balance in accounts receivable (value of the population) = \$300,000
- Risk of incorrect acceptance = 10% (which corresponds to a confidence level of 90%)
- Expected misstatement (error) = \$4,000
- Tolerable misstatement (error) = \$10,000

Determining Sample Size (Louwers, page 802-803)

The Monetary Unit Sampling module can be accessed through the “Analysis” tab in IDEA. Using the ***Analysis>Sample>Monetary Unit>Plan*** function, input the parameters established for the audit of Rice Company’s accounts receivable (NOTE: In the following screen shot, data input by the audit team are circled in red, IDEA processing results are circled in green). In this case, the recorded balance of accounts receivable is the population of interest.

⁸ A white paper “Monetary Unit Sampling Explained” by Kurt B. Johnson provides some detailed guidance on the use of various features of IDEA in monetary unit sampling.

Monetary Unit Sampling - Plan

Total Value of Sampled Population

Use values from database field: RECORDED_AMOUNT

Positive values Negative values Absolute values

Value of the sampled population: 300,000.00

Settings

Confidence level (%): 90.00

Amount Percentage

Tolerable error: 10,000.00

Expected error: 4,000.00

Change basic precision pricing (BPP) from 100% : 100

Approximate sample size:	168
Sampling interval:	1,785.71
Sum of tolerable sample taintings:	224.00 %

You may accept the population at the 90.00% confidence level when no more than 2.240000 total taintings are observed in a sample of size 168.

This is the minimum sample size that will allow you to draw the above conclusion.

Buttons: Estimate, Accept, Print, Cancel, Help

Two options exist with respect to entering the value of the population (“Total Value of Sampled Population” in the screen shot): (1) importing the account values from a database field (which would represent Rice’s subsidiary accounts receivable ledger) or (2) directly entering the value of the population. In this case, uncheck the box “Use values from database field” and manually enter the value of the population (\$300,000). After entering the parameters, select “Estimate” and the appropriate sample size and sampling interval are provided (see items circled in green in the screen shot).

Note that the sample size ($n = 168$) and sampling interval (\$1,785.71) slightly differ from those determined through the use of AICPA sampling tables (as shown in Louwers, pages 802 and 805) ($n = 175$ and sampling interval = \$1,714). The reason for these differences is that AICPA tables assume a slightly different statistical distribution of the account balances and will only negligibly affect the determination of sample size and evaluation of results.

After determining the sample size, selecting “Accept” will retain both the sample size and sampling interval for use in later stages of the sampling process.

Selecting the Sample (Louwers, pages 804-806)

Once the sample size has been determined, the audit team can use IDEA to select sample items for examination. Accessing the **Analysis>Sample>Monetary Unit>Extract** function, the audit team should select “Fixed interval” extraction. There are two options for handling “high values” (items that have a recorded value greater than the sampling interval). These items can be included in the aggregate sample or can be written to a separate database.

Note that the sampling interval (\$1,785.71) has been carried over from the sample planning stage and a random starting point of \$612.96 has been provided by IDEA. (As before, items circled in red in the following screen shot would be input or selected by the audit team; items circled in green are the results of processing by IDEA).

Monetary Unit Sampling - Extract

Extraction type

Fixed interval

Cell selection

High value handling

High values in sample as aggregate

High values in database

High value file name: High Values

OK

Fields

Cancel

Help

Numeric field to sample: RECORDED_AMOUNT

Sample interval: 1,785.71

Random starting point: 612.96

Change high value amount: 1,785.71

By choosing “OK”, IDEA selects sample items from the database of transactions or components using the random start, sampling interval, and sample size. The sample selection would proceed as follows:

1. The first item selected would be the 613rd (rounded) dollar in the population.
2. The second item selected would be the 2399th (rounded) dollar in the population ($\$613 + \$1,786 = \$2,399$).
3. Every 1,786th dollar thereafter (sampling interval) will be selected until a total of 168 dollars (the desired sample size) have been selected for examination.

Measuring Sample Items (Louwers, pages 806-807)

After selecting the sample items, the audit team will then conduct the normal auditing procedures to measure the sample items. Let's assume that three misstatements were identified (see Louwers, Exhibit G.4, page 806):

Account	Recorded Balance	Audited Balance	Difference	Tainting Percentage
10	\$1,000	\$ 900	\$ 100	10%
598	3,840	1,920	1,920	50%
1139	525	420	105	20%

Because Account 598 has a recorded balance (\$3,840) greater than the sampling interval (\$1,785.71), the audit team would not need to project this misstatement to the population.

Evaluating Sample Results (Louwers, pages 807-811)

Once the sample items have been measured and any misstatements identified, two options exist for evaluating the sample results in IDEA. One option is to evaluate misstatements in the accounts with recorded balances less than the sampling interval and to then separately consider the misstatement in Account 598. The other is to create two separate files, one with misstatements in accounts with recorded balances less than the sampling interval and one with accounts having recorded balances greater than the sampling interval. Both methods will yield the same results and conclusion; in this example, we illustrate the use of two separate files in sample evaluation. The files created in this example are as follows:⁹

Recorded values less than sampling interval ["Less than SI"]

ACCOUNT	RECORDED_AMOUNT	AUDITED_VALUE
10	\$1,000	\$ 900
1139	525	420

⁹ In an actual application using IDEA, all 168 items selected for examination would be written to a file, with a column for audited value (AUDITED_VALUE) generated by IDEA. IDEA initially sets the audited value as equal to the recorded value, with the audit team making any modifications to the audited value for misstatements identified during the audit.

Recorded values greater than sampling interval ["Greater than SI"]

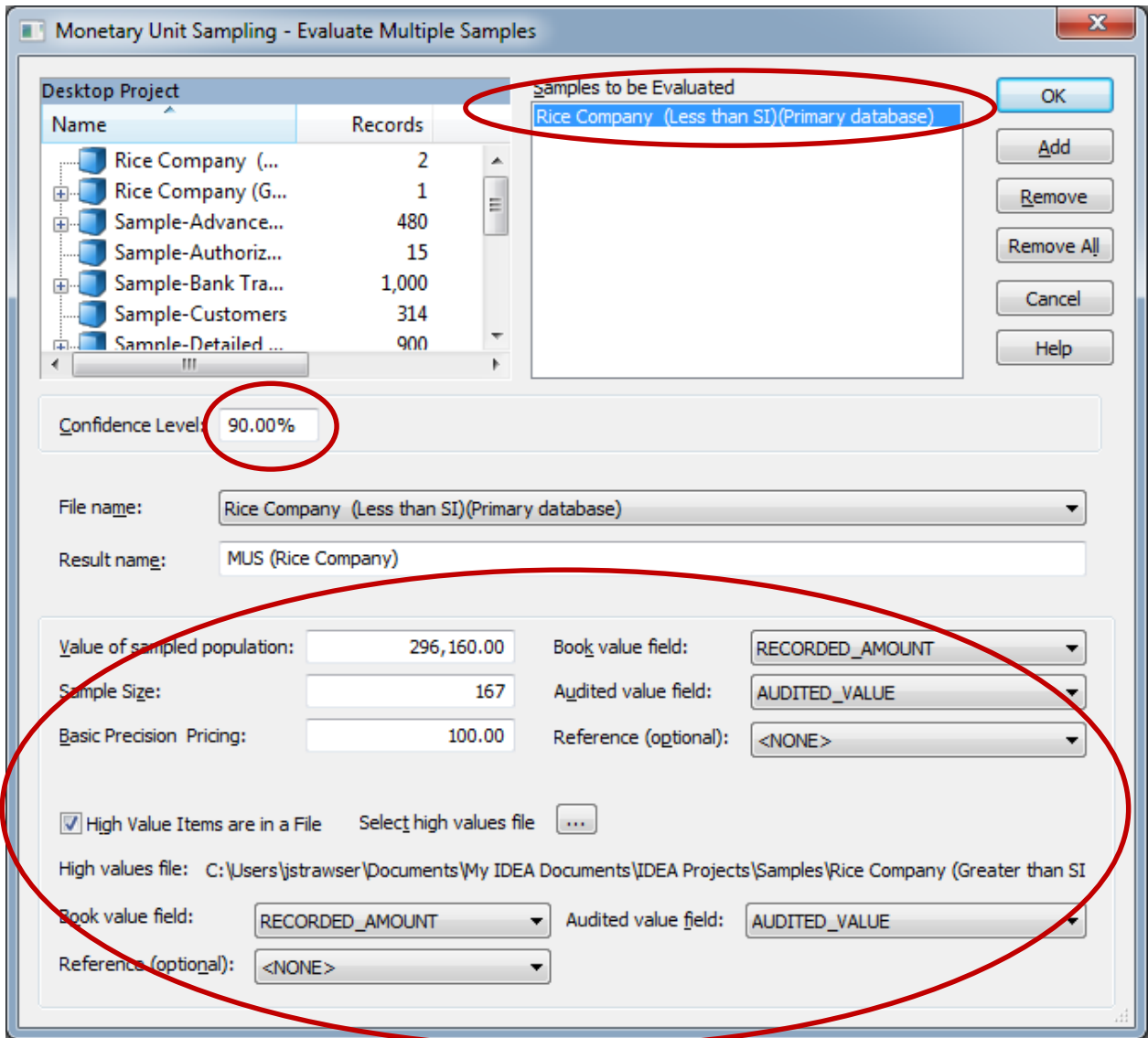
ACCOUNT	RECORDED_AMOUNT	AUDITED_VALUE
598	\$ 3,840	\$ 1,920

To evaluate the sample results, select the **Analysis>Sample>Monetary Unit>Evaluate>** function. The audit team should select the "Multiple Sample" option, as this employs the Stringer Bound methodology assumed by the AICPA sampling process. (This option can be used even in sampling applications in which only a single sample was selected).

Next, the audit team should identify the appropriate files. If MUS was used to select the sample, the file would be automatically populated by IDEA. As noted in the previous section, the audit team has created two files: "Less than SI" and "Greater than SI".

Once the files have been identified or created, the audit team enters the confidence level (90%, which corresponds to a risk of incorrect acceptance of 10%), the value of the population (\$296,160), and the sample size (167). Note that, since high value items are included in a separate file, both the sample size (168) and population (\$300,000) are reduced for the high value item of \$3,840.¹⁰ (Basic Precision Pricing should remain at 100.00). The audit team should also ensure that the Book value field (RECORDED_AMOUNT) and Audited value field (AUDITED_VALUE) are correctly identified for both files. As before, items circled in red in the following screen shot would be input or selected by the audit team.

¹⁰ This assumes that the population contained only one high value item.



At this point, if “OK” is selected, IDEA generates the following output.

MUS - Combined Sample Evaluation Summary

Confidence Level 90.00

Value of the sampled population excluding high values 296,160.00
 Total value of high value items 3,840.00
 Value of the sampled population including high values 300,000.00

Results Excluding High Value Items	Overstatements	Understatements
Combined Sample Size	167	167
Number of errors	2	0
Gross most likely error	532.02	0.00
Net most likely error	-	-532.02
Total Precision	4,329.18	4,054.00
Gross upper error limit	4,861.20	4,054.00
Net upper error limit	4,861.20	3,521.98
Results for High Value Items		
Number of high value items	1	1
Number of errors	1	0
Value of errors	1,920.00	0.00
Results Including High Value Items		
Total number of items examined	168	168
Number of errors	3	0
Gross most likely error	2,452.02	0.00
Net most likely error	2,452.02	-2,452.02
Gross upper error limit	6,781.20	4,054.00
Net upper error limit	6,781.20	1,601.98

Conclusion:
 Based on the combined sample, the most likely total overstatement in the combined population is 2,452.02, and the most likely total understatement is 0.00. You can infer with a confidence of 90.00% that the total overstatement in the combined population does not exceed 6,781.20, and that the total understatement in the combined population does not exceed 4,054.00.

From the above, note the following (refer to the column for “Overstatements”):

- Using a confidence level of 90% (risk of incorrect acceptance of 10%), the audit team examined a population totaling \$300,000. This population had one high dollar value (recorded at \$3,840). (Item #1)
- Under “Results Including High Value Items”, a total of 168 items were examined and 3 errors (all overstatement) were identified.
- The “Gross Most Likely Error” corresponds to the projected misstatement discussed in the text. In this case, the “Gross Most Likely Error” is \$2,452.02 (Item #2) and consists of two components:
 - The tainting percentage multiplied by the sampling interval for non-high value items (\$532.02) (Item #2A)
 - The actual misstatement in the high value item (\$1,920.00) (Item #2B)
- The “Total Precision” of \$4,329.18 (Item #3) represents both the basic allowance for sampling risk and the incremental allowance for sampling risk. Again, these values will differ slightly from those in the text because of the distribution assumed by IDEA.

The “Gross Upper Error Limit” of \$6,781.20 (Item #4) is the sum of the Gross Most Likely Error (Item #2) and Total Precision (Item #3) from above. As noted in the conclusion, the audit team is able to conclude with 90% confidence (representing a 10% risk of incorrect acceptance) that the total overstatement does not exceed \$6,781.20. Since understatement errors are normally not an important

consideration for accounts receivable, the analysis of potential understatement errors (shown in the rightmost column of the output screen shot) is not addressed.

Assignment 14: Comprehensive Monetary Unit Sampling Example (Estimated time 60-120 minutes)

Access the “**Sample-Detailed Sales**” file from IDEA.

1. By examining the file, how many sales transactions are included? Using the “Field Statistics” function, what is the recorded amount of the population of sales transactions after tax (**SALES_PLUS_TAX**)?

2. By double-clicking on the **SALES_PLUS_TAX** column of the file, sort the population by dollar amount. Based on the composition of this population, which specific item(s) might the audit team wish to consider separately as it designs its MUS sampling application?

3. Determine the sample size and sampling interval for each of the following combinations of parameters. Based on comparisons among these scenarios, describe the impact of each of these elements on sample size:

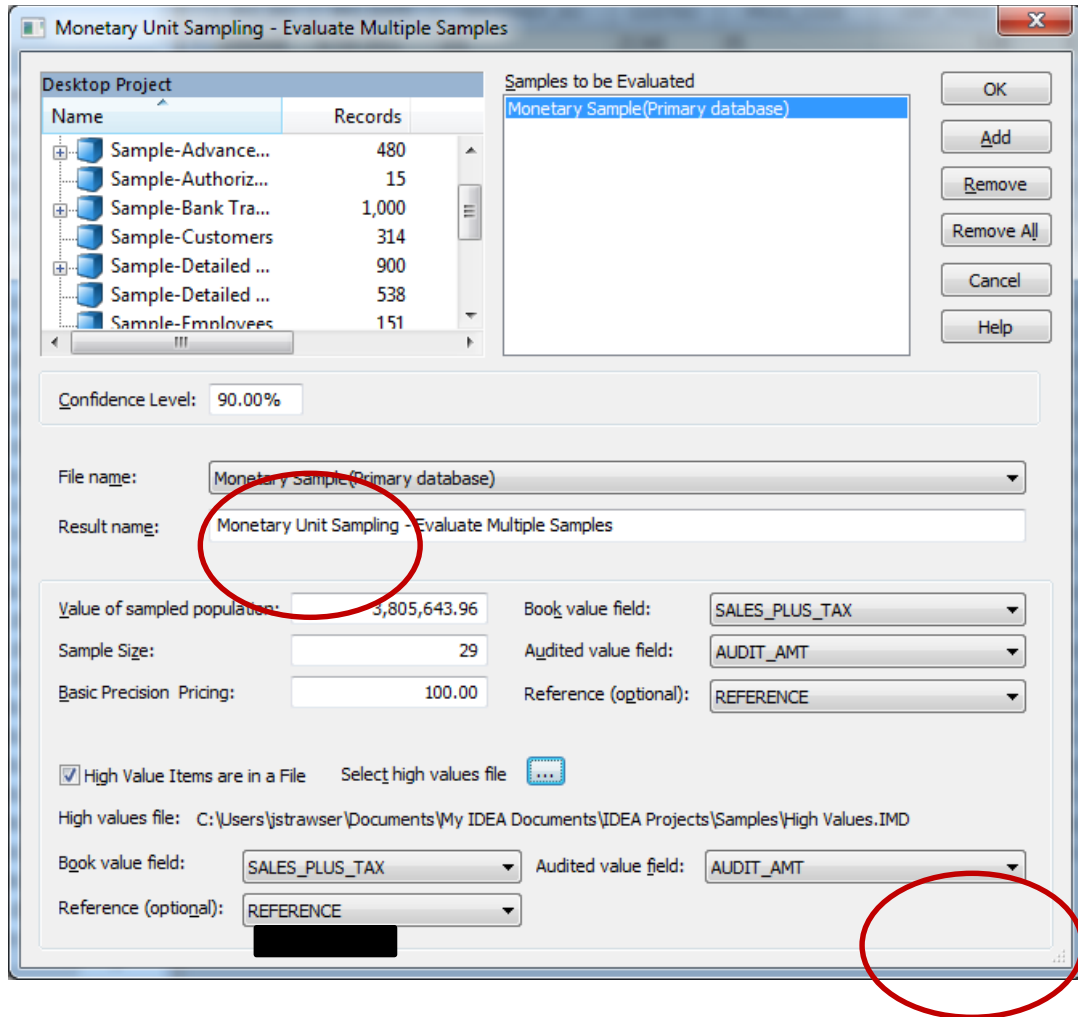
	Confidence (Risk of Incorrect Acceptance)	Tolerable Error (Tolerable Misstatement)	Expected Error (Expected Misstatement)
a.	90% (10%)	\$600,000	\$200,000
b.	90% (10%)	\$1,200,000	\$200,000
c.	90% (10%)	\$600,000	\$100,000
d.	95% (5%)	\$600,000	\$200,000

4. Using the parameters in Scenario (3)(a) above and a random start of \$22,053, select a sample from the population of sales transactions by accessing **Analysis>Sample>Monetary Unit>Extract**. You should include high values in a separate database. In total, how many transactions were selected including both the sample and high values? What percentage of the total number of sales transactions did you select for testing? What percentage of the total dollar value of the transactions did you select? Does this appear to give you adequate coverage of the population? (Use the “Field Statistics” function to summarize the sample items selected.)

5. Assume that the auditor's procedures identified the following misstatements:

Invoice	Recorded Balance	Audited Value
1000544	\$ 4,796.79	\$ 4,612.25
1000612	109,047.95	107,850.00
1000855	139,510.80	135,600.00

Modify the **AUDIT_AMT** column for these three misstatements by accessing the tabs related to the "Monetary Sample" file and "High Values" file (note that one of these misstatements occurred in a transaction that was written to the "High Values" file). After entering the audited values for both files, activate the tab for the "Monetary Sample" file (which will designate this as the primary database – important!) and access **Analysis>Sample>Monetary Unit>Evaluate>Multiple Samples**. While the box is open, specify the high value items in a file as shown below. The screen shot would appear as follows:



After recording these differences in the database and processing, answer the following questions using the IDEA output (the output is shown at the conclusion of this section):

- a. Of the 34 items examined, how many errors were identified?

- b. What is the total gross most likely error?
- c. What is the total gross upper error limit? What accounts for the difference between the gross upper error limit and gross most likely error?
- d. Can the audit team accept the account balance as fairly stated based on the results of this test?

Assignment 14 Solution

1. Using the “Field Statistics” function, identify the number of sales transactions and the recorded amount of the population of sales transactions after tax (***SALES_PLUS_TAX***).

The total recorded amount of the population is \$12,563,283.40, which represents 900 individual sales transactions. Also note that the total value of the positive items is \$12,563,336.11 (there is one negative item of \$52.71).

The screenshot shows a software window titled 'Sample-Detailed Sales' with a 'Properties' pane on the right. The 'Field Statistics' section is active, displaying a table of statistics for the 'SALES_PLUS_TAX' field. The table includes metrics such as Net Value, Absolute Value, # of Records, Positive Value, Negative Value, and various standard deviations and variances.

Numeric Statistics		SALES_PLUS_...
Net Value		12,563,283.40
Absolute Value		12,563,388.82
# of Records		900
# of Zero Items		4
Positive Value		12,563,336.11
Negative Value		-52.71
# of Positive Records		895
# of Negative Records		1
# of Data Errors		0
# of Valid Values		900
Average Value		13,959.20
Minimum Value		-52.71
Maximum Value		4,289,854.50
Record # of Min		73
Record # of Max		188
Sample Std Dev		193,349.57
Sample Variance		37,384,055,345.83
Pop Std Dev		193,242.12
Pop Variance		37,342,517,506.56
Pop Skewness		21.009033
Pop Kurtosis		443.225982

2. By double-clicking on the ***SALES_PLUS_TAX*** column of the file, sort the population by dollar amount. Based on the composition of this population, which specific item(s) might the audit team wish to consider separately as it designs its MUS sampling application?

- The population includes two very large amounts (Invoice 1000091 for \$4,289,854.50 and Invoice 1000457 for \$3,890,106.19) that will likely

include multiple sampling units. These two invoices represent over 65% of the total amount of the population of sales transactions (\$12,563,336.11).

- Four invoices with zero balances (Invoices 1000408, 1000020, 1000885, 1000033) and one invoice (1000775) with a negative balance of \$52.71 require special consideration and may need to be segregated into separate populations, if credit balances or zero balances are important to the audit team’s objectives.

The number of zero items and number of negative items can also be determined by reviewing the “Field Statistics” of **SALES_PLUS_TAX** as in (1) above.

3. Determine the sample size and sampling interval for each of the following combinations of parameters. Based on comparisons among these scenarios, describe the impact of each of these elements on sample size:

	Confidence (Risk of Incorrect Acceptance)	(Risk of Tolerable Misstatement)	Tolerable Error (Tolerable Misstatement)	Expected Error (Expected Misstatement)	Error
a.	90%	(10%)	\$600,000	\$200,000	
b.	90%	(10%)	\$1,200,000	\$200,000	
c.	90%	(10%)	\$600,000	\$100,000	
d.	95%	(5%)	\$600,000	\$200,000	

The screen shot for the parameters in Scenario (a) is shown below. (Note that the sample is selected from the population of positive values, which do not include the negative transaction of \$52.71). Modifying the parameters for confidence, tolerable error, and expected error would result in the following sample sizes and sampling intervals.

Monetary Unit Sampling - Plan

Total Value of Sampled Population

Use values from database field: SALES_PLUS_TAX

Positive values Negative values Absolute values

Value of the sampled population: 12,563,336.11

Settings

Confidence level (%): 90.00

Amount Percentage

Tolerable error: 600,000.00

Expected error: 200,000.00

Change basic precision pricing (BPP) from 100% : 100

Approximate sample size:	96
Sampling interval:	130,868.08
Sum of tolerable sample taintings:	152.83 %

You may accept the population at the 90.00% confidence level when no more than 1.528256 total taintings are observed in a sample of size 96.

This is the minimum sample size that will allow you to draw the above conclusion.

Buttons: Estimate, Accept, Print, Cancel, Help

	Sample Size	Sampling Interval
a.	96	\$130,868.08
b.	32	392,604.25
c.	65	193,282.09
d.	134	93,756.24

Using Scenario (a) as the base case, the following relationships can be observed.

In Scenario (b), the tolerable error increases from \$600,000 to \$1,200,000 while the other parameters are at the same level as Scenario (a). The resultant sample size decreases from 96 items to 32 items. Thus, tolerable error has an inverse relationship with sample size (as tolerable error increases, sample size decreases).

In Scenario (c), the expected error decreases from \$200,000 to \$100,000 while the other parameters are at the same level as Scenario (a). The resultant sample size decreases from 96 items to 65 items. Thus, expected error has a direct relationship with sample size (as expected error decreases, sample size decreases).

In Scenario (d), the risk of incorrect acceptance decreases from 10% to 5% while the other parameters are at the same level as Scenario (a). The resultant sample size increases from 96 items to 134 items. Thus, the risk of incorrect acceptance has an inverse relationship with sample size (as the risk of incorrect acceptance decreases (confidence increases), sample size increases).

4. Using the parameters in Scenario (3)(a) above and a random start of \$22,053, select a sample from the population of sales transactions by accessing **Analysis>Sample>Monetary Unit>Extract**. You should include high values in a separate database. In total, how many transactions were selected including both the sample and high values? What percentage of the total number of sales transactions did you select for testing? What percentage of the total dollar value of the transactions did you select? Does this appear to give you adequate coverage of the population? (Use the "Field Statistics" function to summarize the sample items selected.)

The screen shot for the **Analysis>Sample>Monetary Unit>Extract** function is below. Note that this method uses "fixed interval" extraction and high values are included in a separate database labeled "High Values". The items selected other than those having high values are included in a file named "Monetary Sample". (See circled items in screen shot)

Monetary Unit Sampling - Extract

Extraction type

Fixed interval
 Cell selection

High value handling

High values in sample as aggregate
 High values in database

High value file name: High Values

Numeric field to sample: SALES_PLUS_TAX
Sample interval: 130,868.08

Random starting point: 22,053.00
 Change high value amount: 130,868.08

There are 4 items with a value of 0. Items with a value of 0 will have a 0% chance of being selected.

	Total	Records
<input checked="" type="radio"/> Positive values	12,563,336.11	895
<input type="radio"/> Negative values	(52.71)	1
<input type="radio"/> Absolute values	12,563,388.82	896

File name: Monetary Sample

Create a virtual database

OK
Fields
Cancel
Help

The following 29 items were selected for examination from the population of transactions that were not “high value items” (items defined as exceeding the sampling interval). A number of fields have been hidden to focus on the Invoice number (INV_NO), Invoice date (INV_DATE), Recorded amount (SALES_PLUS_TAX), and Audited value (AUDIT_AMT). The recorded amount of these items is \$861,750.90.

	INV_NO	INV_DATE	SALES_PLUS_TAX	AUDIT_AMT	M
1	1000448	6/19/2011	52.71	52.71	
2	1000780	11/14/2011	1,317.80	1,317.80	
3	1000051	1/27/2011	10,456.74	10,456.74	
4	1000319	3/11/2011	108,718.50	108,718.50	
5	1000410	5/21/2011	40,806.91	40,806.91	
6	1000222	5/30/2011	18,897.25	18,897.25	
7	1000544	10/3/2011	4,796.79	4,796.79	
8	1000781	11/21/2011	91,606.87	91,606.87	
9	1000841	7/28/2011	599.60	599.60	
10	1000414	3/25/2011	52,316.55	52,316.55	
11	1000478	9/12/2011	14,403.55	14,403.55	
12	1000619	10/13/2011	4,282.85	4,282.85	
13	1000405	7/15/2011	69,174.11	69,174.11	
14	1000186	6/10/2011	6,589.00	6,589.00	
15	1000022	7/2/2011	5,374.44	5,374.44	
16	1000856	10/6/2011	58,129.50	58,129.50	
17	1000287	1/15/2011	17,164.35	17,164.35	
18	1000187	2/6/2011	9,224.60	9,224.60	
19	1000690	8/1/2011	13,052.81	13,052.81	
20	1000612	12/5/2011	109,047.95	109,047.95	
21	1000043	7/1/2011	5,529.10	5,529.10	
22	1000029	4/1/2011	3,479.85	3,479.85	
23	1000325	1/20/2011	1,251.91	1,251.91	
24	1000327	4/22/2011	10,423.80	10,423.80	
25	1000083	4/29/2011	78,791.26	78,791.26	
26	1000078	3/11/2011	26,356.00	26,356.00	
27	1000075	5/13/2011	22,264.23	22,264.23	
28	1000750	9/12/2011	58,013.24	58,013.24	
29	1000759	9/16/2011	19,628.63	19,628.63	

The following five items have been identified as “high value items”, as they exceeded the sampling interval of \$130,868.08; the total recorded amount of these items is \$8,757,692.15:

CaseWare IDEA - High Values					
	INV_NO	INV_DATE	SALES_PLUS_TAX	AUDIT_AMT	MUS_REC
1	1000594	11/11/2011	147,573.16	147,573.16	
2	1000609	10/17/2011	290,647.50	290,647.50	
3	1000855	9/26/2011	139,510.80	139,510.80	
4	1000091	3/25/2011	4,289,854.50	4,289,854.50	
5	1000457	7/25/2011	3,890,106.19	3,890,106.19	

Thus, a total of 34 transactions representing sales of \$9,619,443.05 have been selected from a population of 900 transactions that totaled

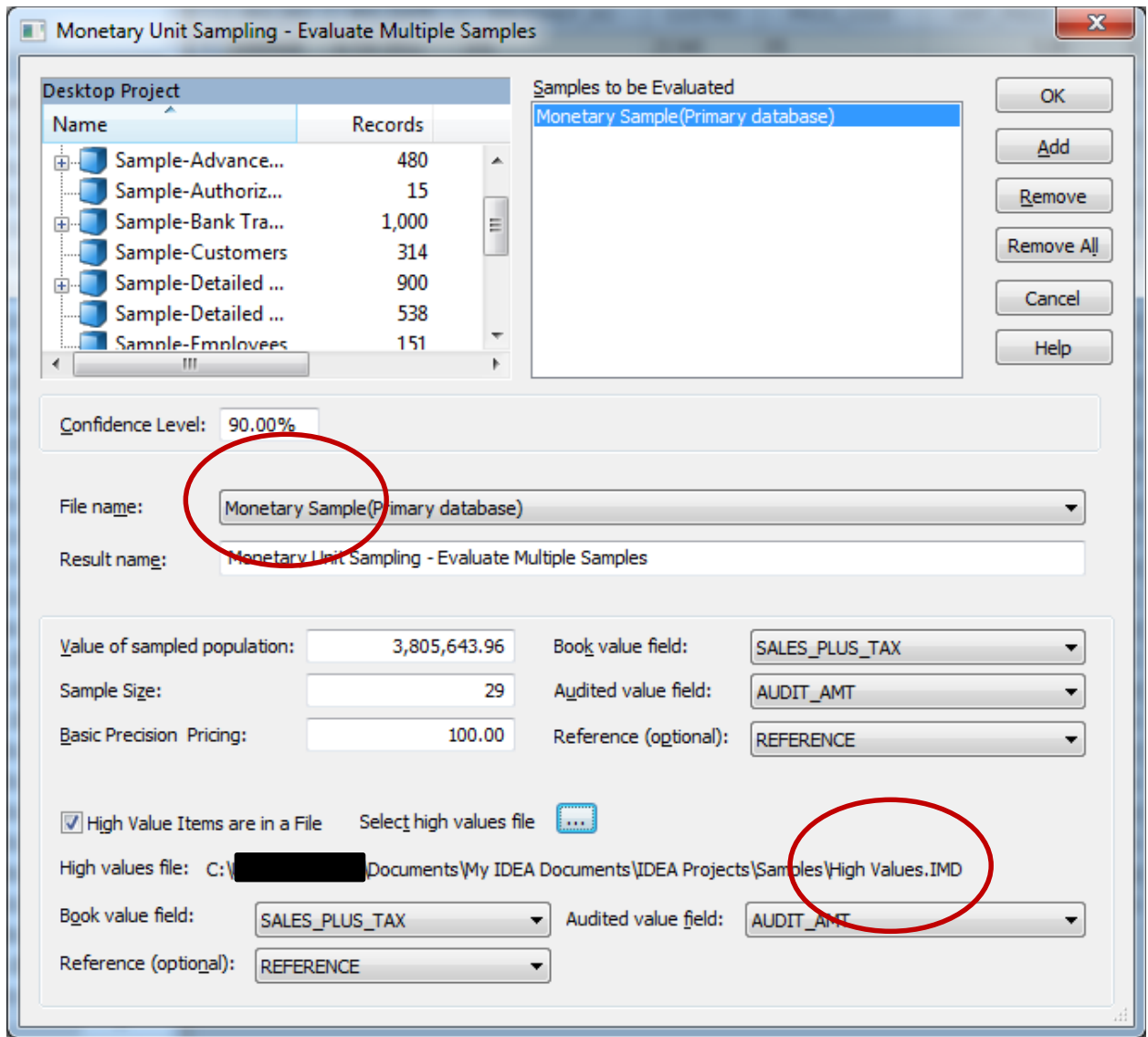
\$12,563,336.11. In sum, MUS allows the auditor to examine 77% of the dollar value of the population ($\$9,619,443.05 \div \$12,563,336.11$) while only evaluating 4% of the transactions ($34 \div 900$).

While the 34 transactions are less than the sample size determined in (3)(a) ($n = 96$), recall that large dollar items may include more than one sampling unit. For example, with a sampling interval of \$130,868.08, the largest transaction (Invoice 1000091 recorded at \$4,289,854.50) would represent at least 32 sampling units ($\$4,289,854.50 \div \$130,868.08$).

5. Assume that the auditor's procedures identified the following misstatements:

Invoice	Recorded Balance	Audited Value
1000544	\$ 4,796.79	\$ 4,612.25
1000612	109,047.95	107,850.00
1000855	139,510.80	135,600.00

Modify the **AUDIT_AMT** column for these three misstatements by accessing the tabs related to the "Monetary Sample" file and "High Values" file (note that Invoice 1000855 is included in the "High Values" file). After entering the audited values for both files, activate the tab for the "Monetary Sample" file (which will designate this as the primary database) and access **Analysis>Sample>Monetary Unit>Evaluate>Multiple Samples**. While the box is open, specify the high value items in a file as shown below.



After recording these differences in the database and processing, answer the following questions using the IDEA output (the output is shown at the conclusion of this section):

- a. Of the 34 items examined, how many errors were identified?

Three (two in “Results Excluding High Value Items” and one in “Results for High Value Items”)

b. What is the total gross most likely error?

The total gross most likely error is \$10,401.01 (see Item #1). Of this, \$6,490.21 relates to projected misstatements and incremental allowance for sampling risk for the sample excluding high value items (Item #2) and \$3,910.80 represents the identified misstatement in the high value items (Item #3).

c. What is the total gross upper error limit? What accounts for the difference between the gross upper error limit and gross most likely error?

The gross upper error limit is \$303,808.48 (see Item #4), comprised of the following:

Total precision (see Item #5)	\$
.....	293,407.47
Gross most likely error (see Item #1)	<u>10,401.01</u>
Gross upper error limit	<u>\$</u>
.....	<u>303,808.48</u>

The difference between the gross upper error limit and the gross most likely error represents the basic precision (\$303,808.48 - \$10,401.01 = \$293,407.47).

e. Can the audit team accept the account balance as fairly stated based on the results of this test?

Because the gross upper error limit (\$303,808.48) is less than the tolerable misstatement (\$600,000), the audit team would accept the account balance as fairly stated.

MUS - Combined Sample Evaluation Summary

Confidence Level 90.00

Value of the sampled population excluding high values 3,805,643.96
 Total value of high value items 8,757,692.15
 Value of the sampled population including high values 12,563,336.11

Results Excluding High Value Items	Overstatements	Understatements
Combined Sample Size	29	29
Number of errors	2	0
Gross most likely error	6,490.21	0.00
Net most likely error	6,490.21	-6,490.21
Total Precision	293,407.47	290,480.00
Gross upper error limit	299,897.68	290,480.00
Net upper error limit	299,897.68	283,989.79
Results for High Value Items		
Number of high value items	5	5
Number of errors	1	0
Value of errors	3,910.80	0.00
Results Including High Value Items		
Total number of items examined	34	34
Number of errors	3	0
Gross most likely error	10,401.01	0.00
Net most likely error	10,401.01	-10,401.01
Gross upper error limit	303,808.48	290,480.00
Net upper error limit	303,808.48	280,078.99

Conclusion:

Based on the combined sample, the most likely total overstatement in the combined population is 10,401.01, and the most likely total understatement is 0.00. You can infer with a confidence of 90.00% that the total overstatement in the combined population does not exceed 303,808.48, and that the total understatement in the combined population does not exceed 290,480.00.



REFERENCES

- IDEA Workbook Version Nine. 2012. Available at <http://www.ideav9.caseware.com/pdf/idea-tutorial.pdf>
- Johnson, K. "Monetary Unit Sampling Explained".
- Louwers, T. R. Ramsay, D. Sinason, J. Strawser, and J. Thibodeau. 2015. *Auditing and Assurance Services* (6th Edition) McGraw-Hill Higher Education, Inc., Burr Ridge, IL;
- PricewaterhouseCoopers. 2015. "Data Driven: What Students Need to Succeed in a Rapidly Changing Business World." Available at: <http://www.pwc.com/us/en/faculty-resource/assets/PwC-Data-driven-paper-Feb2015.pdf>.
- SINTEF. 2013. "Big Data, for better or worse: 90% of world's data generated over last two years." ScienceDaily, May 22.