JFA Journal *of* Forensic *and*
Investigative Accounting
LEARN MORE

Christian's College Fund Store—A Forensic Data Analytics Case Study Using IDEA Software

David C. Hayes
James H. Irving
Michael P. Johnson*

The AACSB's Accreditation Standard A5, Information Technology Skills and Knowledge for Accounting Graduates, requires accredited accounting programs to incorporate technology instruction into their curricula. This standard was established to "develop skills and knowledge related to the integration of information technology in accounting and business" (AACSB, 2018, p. 27). Data analytics is a primary component of Standard A5 and includes statistical techniques, modeling, text analysis, predictive analytics, and visualization. CaseWare IDEA by Audimation Services is an audit analytical and file interrogation software tool designed to help users identify risks, evaluate internal controls, and strengthen fraud detection. The purpose of this paper is to provide another resource for accounting programs and individual accounting faculty members desiring to integrate forensic-related skills and data analytics into their curriculum.

Background

Recent surveys of accounting educators (Dzuranin, Jones, and Olvera, 2018) and accounting practitioners (Lee, Kerler, and Ivancevich, 2018) establish the importance of integrating software tools into the accounting curriculum. These software tools can be divided into several categories, including

- Spreadsheets (e.g., Microsoft Excel)
- Databases (e.g., Microsoft Access)
- Authoritative accounting literature (e.g., FASB Codification, Internal Revenue Code)
- Audit analytical software (e.g., ACL, IDEA)
- Accounting/ERP software (e.g., Peachtree, QuickBooks, SAP)
- Data visualization software (e.g., Tableau)
- Statistical programming languages (e.g., SAS, SPSS, Stata)
- Other programming languages (e.g., Python, R, Visual Basic)

Prior accounting education studies provide accounting faculty members with a variety of instructional resources using audit analytical software. These studies include Richardson and Louwers (2010), Daigle, Daigle, and Lampe (2011), Antcliff, Doren, Harris, and Hayes (2012), Lehmann (2012), Le and Lehmann (2016), Fay and Negangard (2017), and Lee and Nuxoll (2018). This paper contributes to the instructional resources available for audit analytical software, with a focus on developing and enhancing students' forensic-related and data analytics skills.

Learning Objectives

After completing the case, students should become more proficient in their understanding and utilization of the following 17 data analytics skills:

1. Importing data from Excel/Access files
2. Importing data from PDF files
3. Control totals
4. Generating field statistics
5. Criteria
6. Joins
7. Appending fields
8. Direct extractions

*The authors are, respectively, Professor, Associate Professor, and accounting student at James Madison University.

9. Equation editor
10. Extractions text
11. Duplicate detection
12. Gap detection
13. Summations
14. Stratifications
15. Benford's Law
16. Aging analysis
17. Date functions

Prior to attempting the case, students should first become familiar with IDEA software by completing chapters 1 through 4 of the IDEA Data Analytics Workbook.[1] This workbook introduces the IDEA software tool, which is most useful when investigations involve large datasets and many transactions. The workbook assignment is intended to introduce students to a data analytics software and to establish a baseline level of proficiency in data analysis techniques. The workbook consists of an accounts receivable audit, accounts payable audit, fraud investigation, and inventory analysis. Each audit is accompanied by its own set of data files to be analyzed and skills to be learned during the investigations. After completing the workbook, students further enhance their understanding by applying the forensic data analytics skills to the case.

The Case

Christian's College Fund Store assignment

Christian Jones has started college and needs spending money for books and other miscellaneous expenses. He consulted with his dad and they decided to open a store where Christian and his dad go shopping for good deals at auctions and garage sales and then resell the items online. So, Christian opened his small warehouse store (located at 123 James Street) on October 1st. Christian always buys the inventory items for the same cost per item and sells them for the same sales price per item. Here is a partial listing of the items that he buys with his sales price and his cost per item:

Item #	Description	Sales_Price	Cost
1	Wii Video Games	$ 7.00	$ 1.00
2	Xbox Video Games	$ 9.00	$ 1.00
3	Flash Drives	$ 6.00	$ 1.50
4	Vintage GameBoy DS	$ 30.00	$ 8.00
5	Clothes	$ 15.00	$ 2.00
6	Coats	$ 20.00	$ 2.50
7	Toy Race Cars	$ 4.00	$ 0.75
8	Toy Trains	$ 15.00	$ 3.00
9	Vintage Matchbox Cars	$ 8.00	$ 1.00
10	Tables	$ 50.00	$ 7.00

Since dad is interested in helping Christian but is reluctant to give him money, he gave Christian a list of his students to sell to (dad is an educator). A brief description of how Christian's store process works is as follows: Customers can either purchase items through Ebazon (similar to eBay/Amazon) or through Michael's List (similar to Craig's List). Items sold on Ebazon are delivered the same day by a designated shipping clerk (Carl Smith) but are not paid for until later (all sales are on account since they are local customers and we can hunt them down if needed for nonpayment). Those items sold on Michael's List are sold and delivered by Sales Reps/Delivery and also are paid for later (sold on account).

Christian has decided to employ mainly family members for Michael's List sales and a friend to help with online sales. Here is a list of his employees:

---

[1] The IDEA software, workbook, and workbook data files are free downloads, available by completing a request form at the following URL: https://idea.caseware.com/about/contact-us/?form=contact-a-partner. The software does not work well with Macs, but it can be downloaded to computer labs for students who do not have PCs. Version 10 was the current version of the workbook for this case.

Employee_#	L_Name	F_Name	Position
1	Jones	Christian	Owner
2	Jones	Caroline	Sales Rep
3	Jones	David	Sales Rep
4	Jones	Hunter	Sales Rep
5	Smith	Carl	Shipping Clerk
6	Sale	Internet	Computer Sales

Christian uses Microsoft Access to record his sales and has exported his data to an Excel file for you to import and analyze. The sales tab contains the date of the sale, the invoice #, the customer # and the employee # of the employee making the sale. Here are a few examples of sales.

Date	Invoice_#	Customer_#	Employee_#
10/1/2019	1001	51	4
10/1/2019	1002	16	6
10/1/2019	1003	27	6
10/1/2019	1004	71	6
10/1/2019	1005	48	6
10/1/2019	1006	47	6

Using the previous tables, you can see that Invoice #1003 was to Customer #27 and was an online Ebazon sale (Computer Sale = Emp #6). The sales database is organized so that the sales detail of each invoice is in a separate row (sales detail tab). Here are a few examples of invoice details:

Inv_#	Item_#	Qty
1001	6	3
1001	10	6
1001	15	6
1001	19	3
1002	7	25
1002	14	12
1002	16	21
1003	6	15
1003	16	15

The customer on Invoice #1003 bought 15 Coats (item #6) and 15 Coloring Books (item #16). In the Cash Receipts tab, you can see that Customer #27 paid for invoice #1003 on 10/19/2019—18 days later.

Date Paid	Invoice_#
10/18/2019	1362
10/18/2019	1371
10/18/2019	1377
10/19/2019	1003
10/19/2019	1008
10/19/2019	1010

Assume that each invoice is paid in full when recorded as received (i.e., there are no partial payments).

Christian has heard that your team is good at data analytics and wants you to answer the following questions. Since he realizes that you will be billing him based on your time, he would like you to track your time on this assignment. He also

would like a complete list of all the procedures (step-by-step instructions) you performed so he can replicate what you did – screen captures of commands and results are helpful!

For each of the questions, you should provide the procedures you performed, your answers, and the time taken to answer each question. It is strongly suggested that you work as a team to get the data into IDEA and then get organized for the analysis. Maybe even answer one question together. Then, you should copy all the files and each team member should analyze the data individually and then compare with your team members once all have attempted it. Part of your grade will be based on accuracy and part of your grade will be based on documenting the process you used to arrive at an answer.

Required:

- Only use IDEA to analyze the data provided and answer the questions
- Create a typed list of procedures and steps you performed to analyze the data whether you find anything or not. Part of the assignment grade is based on what IDEA skills you apply
- Insert screen captures and descriptions to show how you completed the procedures
- Explain your conclusions and findings based on the procedures you performed—assume that your report is being provided to Christian or the Bank
- You are limited to 20 pages (single-spaced, 12pt Times New Roman font), so be thorough, but concise
- The reports should be for the 4th quarter (October 1, 2019 through December 31, 2019)
- Grading: This assignment is worth 80 points and is due at the beginning of class on the assigned day. One team member should upload the solutions file to the course management system and the file name should include the last names of your team members in alphabetical order

Here is an example of procedures and documentation:

1. Used the data file "Revenue Cycle Data.xls" to import data into IDEA
2. Created a folder for the project: Christian College Fund Store
3. Selected Create and External Project—located in Christian College Fund Store
4. After copying the Revenue Cycle Data file to the source folder, imported all the tabs and indicated that first row was field names—named the output file CCFS
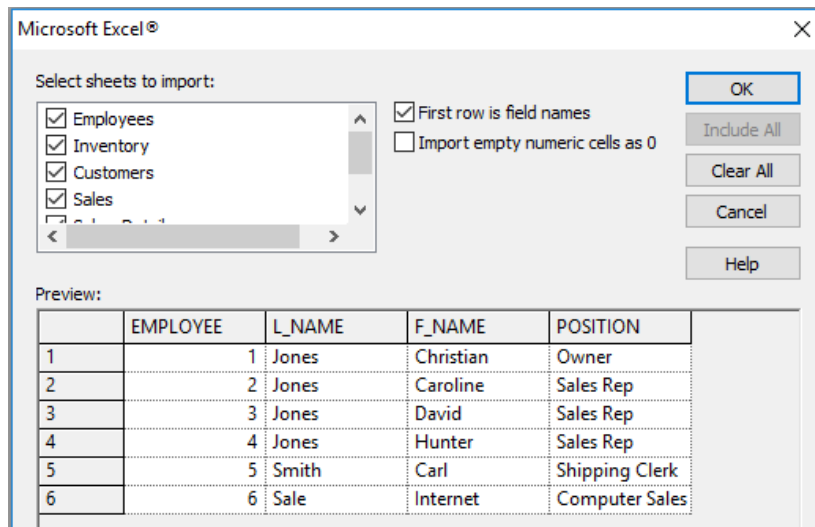
Verified the number of records in IDEA to the spreadsheet Revenue Cycle Data.

Desktop Project	
Name	Records
CCH-Cash_Receipts	1,241
CCH-Customers	72
CCH-Employees	6
CCH-Inventory	20
CCH-Sales	1,331
CCH-Sales_Detail	3,385

Inventory = 20 items, Employees = 6, Customers = 72, etc.

Analyzed the type of data (Numeric, Character, Date) fields to ensure that it is ready for analysis (convert items as needed).

This should be the start of the first page you turn in.

Team # (no names):___

Question #1: Christian wants you to create a set of financial statements for him for the 4th quarter. His financial statements will only include sales, cost of goods sold and gross profit (i.e., do not worry about Other Expenses yet).

Procedures:

Answer:

Sales Revenue	$
Less: Cost of Goods Sold	$_____ .
= Gross Profit	$
Less: Other Expenses	$ 0
= Net Income	$_____

Time to complete this request: ___ Minutes

Question #2: Christian's small warehouse is on James Street and he wants to know which customers live near his warehouse. Which customers live on "Port" Street?

Procedures:

Answer:

Time to complete this request: ___ Minutes

Question #3: Christian wants you to check for missing or duplicate invoices. Per his records, the first invoice is #1001 and the next invoice is #2333.

Procedures:

Answer:

Time to complete this request: ___ Minutes

Question #4: Christian wants to know which (if any) of his customers did not buy anything from him in the 4th quarter. His dad will be unhappy if the customers do not buy things from Christian because then dad may have to give Christian money.

Procedures:

Answer:

Time to complete this request: ___ Minutes

Question #5, Part A: Christian wants you to check for fraud using Benford's law (1st digit only) in the 4th quarter. See if his family (Caroline, David, and Hunter) appear to be working together and committing fraud based on the total invoice amounts.

Procedures:

Answer:

Time to complete this request: ___ Minutes

Question #5, Part B: Since collusion is rare, now check the individual employees (Caroline, David or Hunter) using Benford's first digit analysis.

Procedures:

Answer:

Caroline:

David:

Hunter:

Time to complete this request: ___ Minutes

Question 6: Christian wants you to calculate the amount of cash he has collected in the 4th quarter and the amount of accounts receivable he should have as of 12/31.

Cash Collection Calculation

Procedures:

Answer: $

Time to complete this request: ___ Minutes

Accounts Receivable Calculation

Procedures:

Answer: $

Time to complete this request: ___ Minutes

Question #7: Christian wants you to help him with his business by showing him how much and what percentage of 4th quarter sales revenue Caroline, David, and Hunter have in the $250 to $500 range. That is, what is total 4th quarter dollar sales for each salesperson, and what percentage is this of his/her total 4th quarter sales?

Procedures:

Answer:

Caroline's total 4th quarter sales in the $250 to $500 range are $_____. This amount is _____% of her total 4th quarter sales.

David's total 4th quarter sales in the $250 to $500 range are $_____. This is _____% of his total 4th quarter sales.

Hunter's total 4th quarter sales in the $250 to $500 range are $_____. This is _____% of his total 4th quarter sales.

Which of these three employees has the highest percentage of their individual sales revenue in the $500 to $750 range?

Answer:

_____ had $_____ in the $500 to $750 range. This is _____% of his/her total 4th quarter sales.

Time to complete this request: ___ Minutes

Question #8: Now assume that Christian's business has been growing and he wants to buy a bigger warehouse building and has gone to the local bank for a loan. The bank wants to know how he is doing and he provides the following income statement for the year ended December 31, 2019:

Sales Revenue	$813,768
Less: Cost of Goods Sold	$129,769
= Gross Profit	$683,999
Less: Other Expenses	$583,999
= Net Income	$100,000

Assume you are part of the loan committee and are instructed to investigate Christian's financial statements using the data that he supplied (you still cannot use Excel) to help determine whether or not to loan him money:

Comment on Christian's 4th quarter sales as well as any unusual patterns in revenue recognition (e.g., sales improperly included in or improperly excluded from the 4th quarter).

Procedures:

Answer:

Time to complete this request: ___ Minutes

Question #9: In the past, the bank has lent money to businesses that had too many sales concentrated in one customer. For example, small businesses that only sell to a big company like Wal-Mart are at its mercy should they decide to switch suppliers. Is this a concern for Christian?

Procedures:

Answer:

Time to complete this request: ___ Minutes

Question #10, Part A: Calculate an aging analysis of accounts receivable as of 12/31/2019 based on the following industry guidelines (each bullet point represents an aging "bucket"):

- Current (within 15 days of the invoice date) = 1% chance of being written off. Note that 12/31 is 1 day old
- 1 to 15 days past due = 5% chance of being written off
- 16 to 30 days past due = 25% chance of being written off
- Over 30 days past due = 50% chance of being written off

Calculate the amount of the journal entry Christian will need to record to reflect his store's 2019 bad debt expense by performing the following steps:

- Create an "Age_of_Invoice" field by comparing the invoice date to 1/1/2020
- Create a field listing the aging buckets ("Est_BD_Percent") and have the column populate based on how old the invoice is. Use the percentages provided above
- Create a field to estimate the annual bad debt expense that is expected to be written off ("Est_WO") by multiplying Sales_Sum x Est_BD_Percent
- Show in one report the sum of the number of records for each percentage, the accounts receivable dollar amounts associated with that estimate, and the Est_WO amount for each aging bucket
- Total the report to show the total bad debt expense for the journal entry

Procedures:

Answer:

What is the journal entry to record Christian's 2019 estimate of bad debt expense?

Time to complete this request: ___ Minutes

Question #10, Part B: Now, run two aging reports in IDEA based on the above dates. Use revenue for one of the aging reports and the estimate of bad debt expense for the other aging report. Show screen captures of your results – do they match your calculations from above?

Procedures:

Answer:

Time to complete this request: ___ Minutes

Evidence of Case Efficacy

After completing the workbook, students were asked to complete a pre-test survey rating their perceived proficiency pertaining to 17 IDEA skills on a scale from 0 (No idea how to use the function) to 100 (Expert level of confidence). After completing the case, students were asked to complete a post-test survey updating their perceived proficiency rating on the same 17 IDEA skills. Table 1 reports changes in students' proficiency levels related to the different IDEA skills.[2] Panel A shows that the differences in means between the pre-test and post-test are statistically significant for each of the 17 skills utilized in the case (predicted relation +). As a control, two skills (generating random numbers and pivot tables) in the workbook that were not needed for the case were included in the survey. The differences in means between the pre-test and post-test for these two skills are statistically insignificant lending more support for the findings. Across all 17 skills, students' average proficiency ratings increased 17.5 points (34.3 percent), from an average of 51.0 in the pre-test survey to an average of 68.5 in the post-test survey. Further examining the magnitude of increases in the mean ratings for individual skills, we observe large improvements (greater than 20-point increase in the mean student proficiency ratings between the pre-test and post-test) in the skills related to joins, duplicate detection, gap detection, Benford's Law, and aging analysis. [See Table 1, Panel A, pg. 515]

In Panel B of Table 1, for each of the 17 skills assessed, we eliminate any student who responded in the pre-test survey with a proficiency rating of 90 or above. It is likely that these students have previous training on these skills. Because we wish to gauge the learning of students who are new to learning these skills, we eliminate potential "ceiling effects" by removing these prior experts from the sample. After these adjustments, the differences in means for all 17 skills continue to be statistically significant. Students' average proficiency ratings increased 20.5 points (45.2 percent), from an average of 45.6 in the pre-test survey to an average of 66.1 in the post-test survey. Considerable improvements in individual skills (greater than 20-point increase in the mean student skill level ratings between pre-test and post-test) now include importing PDF files and control totals, in addition to the five skills from Table 1, Panel A (joins, duplicate detection, gap detection, Benford's Law, and aging analysis). [See Table 1, Panel B, pg. 516]

Generally, students agreed the case was effective (80 percent agreed that they learned from the workbook and 83 percent agreed that they learned from the case). Some feedback received from the students included: "I liked combining different databases to achieve the results I wanted", "I liked seeing how we could apply what we have learned into an actual application. It was rewarding once we figured out how to do a particular question", "I liked how relevant this software is to my future in the accounting profession", "It was good to work with software that firms actually use day-to-day".

Concluding Remarks

Given the documented importance of software tools like IDEA (Dzuranin et al., 2018; Lee et al., 2018), this paper contributes to the instructional resources available for audit analytical software. Students complete two assignments – an introductory workbook followed by a case study—using IDEA. Students demonstrated a basic knowledge of key data analytics skills after completing the workbook. However, after completing the case, their knowledge significantly increased across each of the data analytics skills.

---

[2] Pre-test and post-test surveys asked for the maiden name initials of a student's mother and the student's favorite color. These identifiers enabled us to match the pre-test and post-test surveys, yet also preserve student anonymity. The identifiers were not identical in five pre-test and post-test matches (e.g., AEB pink and AEB green). All inferences in Panel A of Table 1 are identical after eliminating these five student nonmatches and re-running the tests for differences in means.

References

Antcliff, M., R. Doren, L. Harris, and D. C. Hayes. 2012. A case to provide students practice in advanced functions of IDEA software. *AIS Educator Journal 7*(1): 69–73.

Association to Advance Collegiate Schools of Business International (AACSB). 2018. 2018 eligibility procedures and accreditation standards for accounting accreditation. Retrieved from https://www.aacsb.edu/-/media/aacsb/docs/accreditation/accounting/standards-and-tables/2018-accounting-standards.ashx.

Daigle, J. J., R. J. Daigle and J. C. Lampe. 2011. Using ACL Scripts to Teach Continuous Auditing/Monitoring: The Tremeg Case. *Journal of Forensic & Investigative Accounting 3*(2): 277–389.

Dzuranin, A. C., J. R. Jones, and R. M. Olvera. 2018. Infusing data analytics into the accounting curriculum: A framework and insights from faculty. *Journal of Accounting Education 43*: 24–39.

Fay, R., and E. M. Negangard. 2017. Manual journal entry testing: Data analytics and the risk of fraud. *Journal of Accounting Education 38*: 37–49.

Le, N., and C. M. Lehmann. 2016. Purchasing process internal control assessment: A comprehensive case study using data analytic software. *AIS Educator Journal 11*(1): 9–15.

Lee, C., and C. Nuxoll. 2018. A Forensic Accounting Case: A Data Analytic Mindset. *Journal of Forensic & Investigative Accounting 10*(3): 436–437.

Lee, L., W. Kerler, and D. Ivancevich. 2018. Beyond Excel: Software tools and the accounting curriculum. *AIS Educator Journal 13*(1): 44–61.

Lehmann, C. M. 2012. Integrating generalized audit software and teaching fraud detection in information systems auditing courses. *Journal of Forensic & Investigative Accounting 4*(1): 319–368.

Richardson, R. C., and T. J. Louwers. 2010. Using computerized audit software to learn statistical sampling: An instructional resource. *Issues in Accounting Education 25*(3): 553–567.

Table 1: Student Self-Assessment of Skill Proficiency Using IDEA

Panel A: All students with a pre-test and post-test match

Skill	Predicted relation	N	Pre-test mean (std. dev)	Post-test mean (std. dev)	Percent change	p-value
Import Excel/Access files	+	35	62.1 (27.4)	75.1 (26.9)	20.9%	0.003***
Import PDF files	+	35	54.2 (26.9)	71.1 (25.5)	31.2%	0.001***
Control totals	+	35	62.4 (29.2)	80.0 (22.7)	28.2%	0.001***
Field statistics	+	35	65.1 (27.7)	78.3 (24.0)	20.3%	0.001***
Criteria	+	35	51.9 (30.0)	66.1 (24.7)	27.4%	0.004***
Joins	+	35	34.4 (22.0)	76.5 (22.5)	122.4%	<0.001***
Appending fields	+	35	70.7 (24.4)	78.5 (22.7)	11.0%	0.043**
Direct extractions	+	35	69.3 (25.2)	75.1 (22.8)	8.4%	0.033**
Equation editor	+	35	66.7 (26.7)	71.2 (24.1)	6.7%	0.076*
Extractions text	+	35	46.6 (27.2)	59.4 (26.6)	27.5%	0.008***
Duplicate detection	+	35	37.3 (24.8)	64.5 (27.7)	72.9%	<0.001***
Gap detection	+	35	30.1 (19.6)	63.9 (28.2)	112.3%	<0.001***
Summations	+	35	45.6 (27.4)	62.2 (26.6)	36.4%	0.002***
Stratifications	+	35	52.2 (24.0)	64.3 (26.2)	23.2%	0.004***
Benford's Law	+	35	37.7 (25.4)	62.2 (27.3)	65.0%	<0.001***
Aging analysis	+	35	38.2 (27.0)	59.1 (30.4)	54.7%	0.001***
Date functions	+	35	42.9 (28.9)	57.2 (27.2)	33.3%	0.001***
Generating random numbers	+/–	35	35.6 (23.0)	35.3 (26.2)	–0.8%	0.939
Pivot tables	+/–	35	41.8 (27.9)	38.6 (31.1)	–7.7%	0.423

Pre-test mean refers to the ratings that students assigned themselves (0-100) after completing the IDEA workbook assignment. Post-test mean refers to the ratings that students assigned themselves after completing the Christian's College Fund Store case assignment. Predicted relation is how the mean was expected to change from the pre-test to the post-test. *, **, *** denotes statistical significance at a probability of <0.10, <0.05, and <0.01, respectively (one-tailed test where a relation is predicted and a two-tailed test otherwise).

Table 1: Student Self-Assessment of Skill Proficiency Using IDEA (continued)

Panel B: Eliminating students with ceiling effects (pre-test ratings ≥90)

Skill	Predicted relation	N	Pre-test mean (std. dev)	Post-test mean (std. dev)	Percent change	p-value
Import Excel/Access files	+	28	53.4 (23.4)	70.2 (27.7)	31.5%	0.002***
Import PDF files	+	32	49.9 (23.9)	69.4 (25.9)	39.1%	<0.001***
Control totals	+	28	53.3 (25.4)	78.8 (23.9)	47.8%	<0.001***
Field statistics	+	26	54.0 (23.1)	73.0 (25.3)	35.2%	<0.001***
Criteria	+	30	44.8 (26.1)	63.6 (22.5)	42.0%	<0.001***
Joins	+	34	32.7 (20.1)	76.7 (22.8)	134.6%	<0.001***
Appending fields	+	25	60.4 (21.1)	76.2 (21.6)	26.2%	0.001***
Direct extractions	+	26	59.4 (21.5)	69.9 (23.9)	17.7%	0.003***
Equation editor	+	26	56.4 (23.1)	63.4 (22.8)	12.4%	0.045**
Extractions text	+	32	42.1 (23.6)	57.5 (27.0)	36.6%	0.003***
Duplicates detection	+	34	35.5 (22.6)	63.4 (27.4)	78.6%	<0.001***
Gap detection	+	35	30.1 (19.6)	63.9 (28.2)	112.3%	<0.001***
Summations	+	34	44.0 (26.1)	61.6 (26.7)	40.0%	0.001***
Stratifications	+	34	50.8 (22.9)	63.4 (26.1)	24.8%	0.003***
Benford's Law	+	33	33.9 (20.6)	60.7 (27.5)	79.1%	<0.001***
Aging analysis	+	33	34.4 (22.8)	57.4 (30.4)	66.9%	<0.001***
Date functions	+	33	39.4 (25.8)	55.1 (26.5)	39.8%	0.001***
Generating random numbers	+/−	34	33.8 (20.9)	33.9 (25.4)	0.3%	0.978
Pivot tables	+/−	32	36.8 (23.5)	33.4 (27.2)	−9.2%	0.445

Pre-test mean refers to the ratings that students assigned themselves (0-100) after completing the IDEA workbook assignment. Post-test mean refers to the ratings that students assigned themselves after completing the Christian's College Fund Store case assignment. Predicted relation is how the mean was expected to change from the pre-test to the post-test. *, **, *** denotes statistical significance at a probability of <0.10, <0.05, and <0.01, respectively (one-tailed test where a relation is predicted and a two-tailed test otherwise).