Analyzing Data for Decision Making: Integrating Spreadsheet Modeling and Database Querying

A. Faye Borthick
Georgia State University

Gary P. Schneider
California State University, Monterey Bay

Therese R. Viscelli
Auburn University

ABSTRACT: Although students begin with a spreadsheet to analyze the effect of a change in credit score cutoff on loan performance on sales, they discover that this approach becomes unwieldy as the analysis becomes more complex. As more attributes from more sheets are required, the spreadsheet solution becomes difficult to implement or audit. The case thus motivates students to develop the database skills of importing the spreadsheet data into a database manager, joining tables, and performing the analysis using database querying. The case requires students to think critically to model the business situation in spreadsheet formulas and database query expressions, structure relationships among the attributes across data tables, and manipulate attributes to achieve the business objective. The case is suitable for courses in accounting systems, managerial accounting, decision making, and database systems. Objective questions are provided for assessing students' ability to analyze transaction records using both spreadsheet and database tools.

Keywords: business analytics; credit score analysis; critical thinking; data analytics; data reporting; decision making; querying; spreadsheet model.

THE CASE

The Scene

Executives at YouDriveOut, a chain of used car lots, are brainstorming about how to ensure better financial performance. Alicia is the chief financial officer (CFO), and Manuel is the controller.

Alicia: Manuel, I'm thinking about switching our minimum credit score cutoff from 430 to 450 to improve our overall loan performance, that is, to increase the proportion of payments that our customers make on loans.

Manuel: You want me to compare the loan default rates for these?

Alicia: Yes, that's the idea.

Manuel: Any particular format?

Alicia: No, but whatever you create should show the difference in overall loan currency with the impact on sales for the two cutoff numbers. And be sure to do the analysis by dealer with the dealer street address. Pivot tables could be useful. It would also be helpful to see the analysis by vehicle category.

Requirements

1. Plan the analysis by completing the following steps based on the conversation:
   a. State the overall objectives of the analysis.
   b. Identify assumptions embedded in the conversation.
   c. Verify that attributes required for the analysis are available in creditScore.xlsx. Data definitions for the attributes are available in Table 1, and a note about primary keys and data types in Microsoft Excel® and Access® is available in Appendix A.
   d. Based on credit score values, which set of loans should be compared to which other set of loans?
   e. Decide the use of each attribute in the analysis. That is, how is each attribute relevant to the analysis?
   f. Scan the data values to find outliers or other situations at odds with assumptions embedded in the conversation.
   g. Specify formulas (spreadsheet) or expressions (database manager) for new variables to be constructed from the attributes supplied in the file creditScore.xlsx.
      To find web-hosted explanations of specific functions and operations, search the web. In each search, include the version number of the program, for example:
         Excel 2013 pivot tables
   h. If you had your choice of spreadsheet modeling or database querying, which would you choose and why?
2. Use your responses in the first requirement to guide an analysis that compares the loan performance for the two cutoff points for the loans in the file (creditScore.xlsx) and provide the executives the results in a form that makes sense to them. Use 9/1/2013 as the reporting date. Perform the analysis as follows:
   a. Develop a spreadsheet model for the analysis.
   b. Import the data (loan, dealer, sales, and acquisitions sheets) into a database manager.
   c. Query to analyze the data in the database manager.
3. Write a one-paragraph memo to the CFO with your recommendation and rationale for it.
4. Based on your experience in the case:
   a. Characterize criteria for choosing analysis tools.
   b. Identify audit and control issues associated with spreadsheets and databases.
5. Answer questions your instructor will provide about the business problem and its analysis.
6. Optional: In addition to the analysis performed on the supplied data, what other analyses would be relevant to YouDriveOut management in deciding whether to change the credit score cutoff?

TABLE 1

Data Attributes
creditScore.xlsx

Sheet/Attribute[a]	Explanation
loan: Loans that the company has made to customers buying its used vehicles	
loanID	Unique identifier for a loan
dealerID	Unique identifier for a dealer
loanAmount	Dollar amount of a loan extended to a customer for a vehicle
downPayment	Dollar amount of the down payment a customer made for a vehicle. The sum of loanAmount + downPayment = the sale price of a vehicle sold on loan.
loanDate	Date a loan was initiated
monthlyPayment	Dollar amount of the monthly payment from a customer to repay (satisfy) a loan, including principal and interest
numberPayments	Number of monthly payments required for a customer to pay off a loan
paymentsMade	Number of monthly payments a customer has made as of the reporting date
creditScore	A score created by a credit-rating agency, e.g., Equifax, Transunion, or Experian, representing the credit worthiness of an individual. A typical range is 300-850 where the lower the score, the less credit worthy the individual is deemed to be.
dealer: Dealers that sell vehicles for the company	
dealerID	Unique identifier for a dealer
name	Name of the dealer
street	Street address of the dealer
sales: Sales of vehicles	
vehicleID	Unique identifier for a vehicle
saleDate	Date a vehicle was sold
saleDealerID	Unique identifier for the dealer selling a vehicle
loanID	Unique identifier for a loan associated with the sale of a vehicle
acquisitions: Vehicles that the company acquired, usually by purchasing them at auction	
vehicleID	Unique identifier for a vehicle
acquireDate	Date the company acquired a vehicle
category	Category of a vehicle: hybrid, sedan, SUV, or truck
cost	Acquisition cost of a vehicle
make	Brand of a vehicle
model	Model of a vehicle
year	Model year of a vehicle

[a] Sheet names appear in bold.

APPENDIX A

Note about Primary Keys and Data Types in Microsoft Excel® and Access®

In Access®, each table has a primary key that uniquely identifies each row. The best practice is to designate an attribute that makes business sense, e.g., loanID uniquely identifies a loan, distinguishing it from all other loans in the table. If no attribute makes business sense, Access® will generate a primary key consisting of a sequence number beginning with 1.

Number representation matters because different representations may give different results in the far-right decimal digits, a condition that is sure to befuddle naïve users. Getting different results from the same calculations in different programs is not a comforting outcome. Microsoft Excel® represents numeric values as double-precision floating-point numbers, which gives 15 digits of precision with an 11-bit exponent and a 52-bit mantissa. Notwithstanding this level of precision, the representation is generally inexact, which leads to round-off and truncation errors in calculations, including with financial data.

The default data type when numeric data are imported from Microsoft Excel® into Microsoft Access® is double-precision floating point. This representation is undesirable in Access® because of the imprecision of the values and the unnecessary use of

disk space. Each double-precision floating point value takes up 8 bytes of disk space in Access®. To avoid using double-precision floating-point representation when importing data from Excel® to Access®, reset the data type in Access® when the data are exact, e.g., dealerID, loanID, and dollar amount attributes in this case. Here are the integer data type possibilities with the ranges of values they can represent:

Microsoft Access® Integer Data Type (Length)	Range of Representation
Byte (1 byte)	0 to 255
Integer (2 bytes)	−32,768 to 32,768
Long integer (4 bytes)	−2 billion to 2 billion

CASE LEARNING OBJECTIVES AND IMPLEMENTATION GUIDANCE

This case prompts students to make sense of a business situation and analyze transaction data to provide insight for the business to guide action, as organizations are doing (Chen, Chiang, and Storey 2012; McAfee and Brynjolfsson 2012; Davenport 2013), an approach known as data-driven decision making (Redman 2008; Dearborn 2015). Not only will accountants encounter data analytics applied to corporate strategy, but they will also see them applied in accounting and finance. For example, in addition to supporting business strategy, chief financial officers need to use information to meet their regulatory requirements (Gnatovich 2007; Johnson 2012), perform finance functions (Griffin and Danson 2012; Cokins 2013), and detect enterprise fraud (Griffin 2012). PricewaterhouseCoopers (PwC 2015) has offered a white paper on the implications for accounting work in a data-driven business environment, and the business press has reported significant changes in accounting work, shifting to fewer accountants doing more analytic tasks (Johnson 2015; Mims 2015; Monga 2015).

Although the analysis in the case could be performed completely with a database query manager, practicing accountants often begin their analyses with spreadsheets because they are skilled with spreadsheets and often receive data in spreadsheet files. Students who have work experience, including internships, have seen the pervasive use of spreadsheets in accounting. In fact, the familiarity is so strong that many students believe that a spreadsheet program is the only analysis tool they will need in practice. Thus, this case is designed to prompt students to reevaluate that belief and recognize the need to become skilled with both spreadsheet modeling and database querying as analysis tools. In this case, when the spreadsheet model becomes unwieldy for integrating needed data attributes, students import the spreadsheet data into a database query manager to perform the analysis. Furthermore, as spreadsheet solutions become more complex, the likelihood of design errors that are difficult to detect increases (Panko and Aurigemma 2010). A database querying approach is often more desirable in complex situations because all the data manipulations are documented in the queries, making them easier to audit.

From the conversation, students build mental situation models that enable them to make inferences about the business process, an experience that helps them develop situation models for the next business situation they encounter (Zwaan and Radvansky 1998; Barsalou 1999).

The specific learning objectives of the case are for students to learn to:

1. Plan data analysis based on the business situation and the available data.
2. Analyze the data to compare loan performance for the two cutoff points for the credit scores:
    a. Develop a spreadsheet model for the analysis.
    b. Import the spreadsheet data into a database manager.
    c. Develop the analysis with database querying.
3. Write a one-paragraph memorandum to the CFO with a recommendation to raise the minimum credit score or to keep the current minimum score and a rationale for the recommendation.
4. Based on their experience in the case, they learn to:
    a. Characterize criteria for choosing analysis tools.
    b. Identify audit and control issues that can arise when working with spreadsheets and databases.

In Borthick's (1996) categorization of learning objectives for the accounting information systems (AIS) course, the case requires (1) extracting or creating needed information and (2) analyzing information to solve business problems. These objectives incorporate the elements of thought for critical thinking established by Paul (1990), Elder (1999), and Nosich (2011). The terms *critical thinking* and *analytical thinking* are often used interchangeably although *critical thinking* is more commonly used in non-computational disciplines than *analytical thinking*.

Level of Difficulty

The case is intended for students who can perform the spreadsheet operations of creating formulas with IF statements, VLOOKUP statements, and date built-in functions, generating pivot tables, formatting attributes, and naming, saving, and retrieving files. The required database query operations are importing data, joining tables, developing expressions, using built-in functions, applying Group By, formatting and sorting results, and naming, saving, and retrieving queries.

Implementation Approaches

Although the requirements direct students to prepare the spreadsheet version first and then do the analysis with database querying, students who already have database querying expertise could be permitted to choose which software tool to use first, based on their understanding of the analysis requirements. The spreadsheet-first approach works well for students with little or no database querying expertise. When they reach the point where the limitations of the spreadsheet approach become apparent, they are receptive to pausing their work long enough to develop sufficient database querying expertise to analyze the data with

the query manager. When our students have not had database querying experience, we have employed this pause-and-develop approach, i.e., they pause the analysis long enough to develop querying skill. To develop database querying skill, we have used Borthick, Jones, and Kim's (2001) case on assuring compliance for responses to website referrals, which requires about ten hours. Bradford, Samuels, and Wood's (2008) Beta Auto Dealers case could also be used for this purpose. Students who are unfamiliar with pivot tables need about an hour to develop proficiency with them.

The second approach, allowing students to choose the software tool, is ideal for students who are already proficient with both spreadsheet modeling and database querying. Posing the software choice to students gives them an opportunity to match the tool to the task and to experience how well the tool matches the task.

Use in Courses

The case is suitable for courses in AIS, managerial decision making, and database systems. It has been used in over 20 sections over three years in undergraduate AIS courses at two large public universities by three instructors. It has been used at a third large public university by a fourth instructor in a capstone course in a master's accounting degree program. It can be completed individually or in teams, in class or outside of class, or in a combination of the two. In our experience, completion time ranges from 4–8 hours for students once they have the needed spreadsheet and database query skills. We make the case materials available to students on a course website.

Students have answered the practice questions in university-provided learning management systems (LMS) that time and score the questions and enable students to see the feedback for each question. The practice questions, which can be offered at any time during students' work on the case, illustrate the kind of thinking students need to do to perform the analysis and give students practice for quiz and exam questions if they are used.

Efficacy

Practitioner Reaction

Fourteen practitioners (partners and other high-level professionals from Big 4 and regional accounting firms and finance-accounting professionals from corporations) reacted to the case in a group setting. They thought it would be a superb learning experience for solving business problems on the job because it captured major features of their work environments. They saw the case as authentic because the assignment was given in a brief conversation with minimal instruction, it concerned a real problem, and it required analyzing data. They confirmed that their environments were characterized by messy data that did not uniformly comply with organizational policies. They were unanimous in wanting to hire graduates who could accept an assignment in a short conversation, get relevant data, analyze the data, recognize when the data did not match everyone's assumptions, accommodate mismatches in the analysis, and report analysis results succinctly with a recommendation and a rationale for it. Because they believed that developing data analytic expertise was a developmental process, they recommended that all accounting courses include authentic learning experiences like this one that require students to solve business problems by analyzing data with software tools.

Faculty Reaction

One professor's reaction to the case, upon first seeing it, was:

I have to use this case in both my AIS and my IS courses. I will no longer have to spend time trying (often unsuccessfully) to get students to understand the differences between Excel and Access and actually have them appreciate the power and practical way you can use a relational database . . . I like the idea of equipping accountants with the tools to analyze large datasets. Further, I like that the case seeks to show the shortcomings of traditional spreadsheet models.

The case asks students to analyze loan data to support a recommendation to either raise the minimum cutoff score for loans extended to customers buying used cars or leave the credit score cutoff unchanged. In our first uses of the case, we observed students being able to think through the logic of what to compare (loan completion rates of loans with credit scores in a stratum at the existing lower cutoff with the next higher stratum) but having difficulty instantiating their logic in a spreadsheet model. The logic required designating the relevant strata of credit scores, calculating the currency of payments for each loan, and calculating the percent of the down payment for each loan from the down payment and the loan amount. Although the students had used and developed spreadsheets in several courses before this one, they apparently overestimated their facility with spreadsheet model building. Students were much more successful in instantiating the logic after we added Requirement 1.a.– 1.g., which prompts them to plan the analysis in a productive way.

Many of our students begin the course believing that spreadsheet models are the only analysis tool that they need. Before we started using this case, many students were skeptical of the need to develop database querying skills. We found that students were more receptive to developing database querying skill after having experienced the need for it in a learning experience such as this case. It also helped that the term *business data analytics* was appearing in the business press.

Student Reaction

Students have had a wide range of reactions to the case. Some were glad they could learn to make sense of a business situation and develop the spreadsheet and querying expertise to offer insight for this kind of problem. Others were sure that the case was not accounting because it did not focus on learning to perform standard accounting calculations from a textbook. In contrast, students who had been on internships before taking the course wished they had taken the course first so that they could have been more productive on the job with data analysis tasks.

The last resistance to learning to query a database disappeared as students created the VLOOKUPs required to match *category* values across the sheets: one VLOOKUP to link *category* from the acquisition sheet to the sales sheet and one VLOOKUP to link *category* from the sales sheet to the setup sheet, where all the other attributes are available. This situation prompted a discussion of the appropriateness of a spreadsheet model for all analysis situations. Students agreed that a way was needed to link the loan data attributes with attributes in the other sheets. At this point, students continued the analysis with a database query manager that could perform this function with a single command and began to develop database query skill for this purpose. To develop the necessary database query expertise for analyzing the loan data by dealer, students worked through Borthick et al.'s (2001) case on assuring compliance for responses to Web site referrals.

Students were relieved that database tables could be joined so easily without the tedious steps involved in making spreadsheet VLOOKUPS and pivot tables do the categorization. Showing how cumbersome a spreadsheet model could be compared to the ease of using a different software tool intrigued students by the idea of developing expertise with the other tool—a database query manager. The idea that spreadsheets and databases could be used together through importing/exporting data motivated them to enhance their skills with both the spreadsheet and the database manager.

In the first year of case use, few of the students had experience developing pivot tables, which are ideal for showing the results by cutoff score strata. Notwithstanding their lack of experience with pivot tables, these students seemed eager to learn how to use them in analyzing the credit score cutoff decision, probably due to having heard about them from internship co-workers and recruiters. By the third year of case use, we observed very few students admitting to not being familiar with pivot tables.

Case Use Where Students Developed Spreadsheets/Queries

One of the sections that used the case was a capstone course in a master of accountancy program at a different large public university. Even though these students had completed an AIS course that included database querying, they all chose to perform the analysis with a spreadsheet. These graduate students had just as much difficulty with the mathematical modeling as the undergraduate AIS students exhibited, that is, they needed substantial help with the mathematical modeling required to create the new attributes to represent credit score strata, loan currency, and proportion paid down. In post-case interviews, the capstone course students reported positive experiences with four-person groups in constructing computational models for the business problem and working with other group members to develop the models.

The difficulty with mathematical modeling experienced by students at both undergraduate and graduate levels suggests that accounting programs should give all students more experience in modeling business situations so that they can contribute to the analyses that are increasingly becoming the basis for data-driven business decisions. An educational diet of memorizing accounting procedures does not appear to prepare students for analyzing accounting data in situations for which no standard accounting computation or schedule applies.

SUMMARY

The contribution of this case is as a learning experience for developing data analytics expertise with both spreadsheet and database query software and for deciding how to use the two tools together to minimize the overall analysis effort. The case is appropriate for students who have mastered basic spreadsheet modeling and database querying skills. Students begin the case by developing analysis objectives from a brief conversation that reveals the business situation and then developing a spreadsheet model to analyze transactions that represent customers' loan payment history. When they attempt to integrate dealer and vehicle category attributes into the analysis, students discover that they need to join data tables, a task for which their spreadsheet models are unwieldy. If they do not have sufficient database querying skill to join tables of loan history and vehicle and dealer attributes and continue the analysis, then they pause the analysis long enough to develop the needed querying skill.

This case enables students to experience, in an authentic way, the need for two approaches to data analysis: spreadsheet modeling and database querying. The use of both approaches gives them the ability to develop criteria for deciding which tool to use for which analyses, including migrating data between the two tools to take advantage of the distinct capabilities of each.

TEACHING NOTES AND STUDENT VERSION OF THE CASE

Teaching Notes and the Student Version of the Case are available only to non-student-member subscribers to *Issues in Accounting Education* through the American Accounting Association's electronic publications system at http://aaapubs.org/. Non-student-member subscribers should use their usernames and passwords for entry into the system where the Teaching Notes can be reviewed and printed. The "Student Version of the Case" is available as a supplemental file that is posted with the Teaching Notes. Please do not make the Teaching Notes available to students or post them on websites.

If you are a non-student-member of AAA with a subscription to *Issues in Accounting Education* and have any trouble accessing this material, please contact the AAA headquarters office at info@aaahq.org or (941) 921-7747.

REFERENCES

Barsalou, L. W. 1999. Language comprehension: Archival memory or preparation for situated action? *Discourse Processes* 28 (1): 61–80.

Borthick, A. F. 1996. Helping accountants learn to get the information managers want: The role of the accounting information systems course. *Journal of Information Systems* 10 (2): 75–85.

Borthick, A. F., D. R. Jones, and R. Kim. 2001. Developing database query proficiency: Assuring compliance for responses to website referrals. *Journal of Information Systems* 15 (1): 35–56.

Bradford, M., J. A. Samuels, and R. E. Wood. 2008. Beta Auto Dealers: Integrating disparate data to solve management problems. *Issues in Accounting Education* 23 (2): 309–326.

Chen, H., R. H. L. Chiang, and V. C. Storey. 2012. Business intelligence and analytics: From Big Data to big impact. *MIS Quarterly* 36 (4): 1165–1188.

Cokins, G. 2013. Driving acceptance and adoption of business analytics. *The Journal of Corporate Accounting & Finance* 24 (2): 69–74.

Davenport, T. H., ed. 2013. *Enterprise Analytics: Optimize Performance, Process, and Decisions Through Big Data*. Upper Saddle River, NJ: FT Press.

Dearborn, J. 2015. *Data Driven: How Performance Analytics Delivers Extraordinary Sales Results*. Hoboken, NJ: Wiley.

Elder, E. 1999. *Critical Thinking: Basic Theory and Instructional Structures Handbook*. Tomales, CA: Foundation for Critical Thinking.

Gnatovich, R. 2007. Making a case for business analytics. *Strategic Finance* (February): 47–51.

Griffin, J., and F. Danson. 2012. Analytics and the cloud—The future is here. *Financial Executive* 28 (9): 97–98.

Griffin, R. 2012. Using Big Data to combat enterprise fraud. *Financial Executive* (December): 45–47.

Johnson, J. E. 2012. BIG DATA + big analytics = big opportunity. *Financial Executive* (July/August): 50–53.

Johnson, K. S. 2015. Outdated: The plain-vanilla accountant. *Wall Street Journal* (May 19): B7.

McAfee, A., and E. Brynjolfsson. 2012. Big Data. *Harvard Business Review* (October): 59–68.

Mims, C. 2015. Data is now the new middle manager. *Wall Street Journal* (April 20): B1, B2.

Monga, V. 2015. The new bookkeeper is a robot. *Wall Street Journal* (May 5): B1, B7.

Nosich, G. M. 2011. *Learning to Think Things Through: A Guide to Critical Thinking Across the Curriculum*. Upper Saddle River, NJ: Prentice Hall.

Panko, R. R., and S. Aurigemma. 2010. Revising the Panko-Halverson taxonomy of spreadsheet errors. *Decision Support Systems* 49 (2): 235–244.

Paul, R. 1990. *Critical Thinking: What Every Person Needs to Survive in a Rapidly Changing World*. Tomales, CA: Foundation for Critical Thinking.

PricewaterhouseCoopers (PwC). 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

Redman, T. C. 2008. *Data Driven: Profiting from Your Most Important Business Asset*. Boston, MA: Harvard Business Review Press.

Zwaan, R. A., and G. A. Radvansky. 1998. Situation models in language comprehension and memory. *Psychological Bulletin* 123 (2): 162–185.