Journal of Accounting Education 64 (2023) 100858

Contents lists available at ScienceDirect

Journal of Accounting Education

journal homepage: www.elsevier.com/locate/jaccedu

A mental model approach to teaching database querying skills with SQL and Alteryx

Lorraine Lee *, Gretchen Casterella

University of North Carolina Wilmington, United States

ARTICLE INFO

Article history:
Received 13 June 2022
Received in revised form 23 June 2023
Accepted 24 June 2023
Available online 30 June 2023

Keywords:
SQL
Data analytics
Mental models

ABSTRACT

This paper describes an instructional approach and exercises to introduce a data analytics tool into an existing course that covers relational databases and structured query language (SQL). SQL and relational databases are traditional topics in many AIS classes and are frequently taught with established technologies such as Microsoft Access. With the onset of newer data analytics tools, educators are faced with the challenge of incorporating emerging technologies while still covering basic foundational concepts. In this paper, we position SQL as the mental model for learning the data preparation and data blending features of a popular analytics tool, Alteryx. We split the traditional SQL content from our course into four increments that increase in complexity. For each increment, we use a multi-step approach to first develop students' mental model of the SQL query code and then apply the mental model to the new environment of Alteryx. For each increment, students solve the same set of problems in SQL and Alteryx, so that the understanding of SQL eases the transition to Alteryx, and the practice in Alteryx reinforces the understanding of SQL.

1. Introduction

Accounting faculty face the ongoing challenge of how to incorporate emerging technologies and skills (e.g., data analytics, robotic process automation, blockchain) into their courses without sacrificing the core knowledge and skills that prepare students for accounting careers and the CPA Exam. The AACSB, in Standard A5, recognizes the importance of information technology for accounting students and requires accounting programs to provide learning experiences that encourage students to master current technology and adapt to emerging technology. Accounting graduates need to "understand the capabilities of these [emerging technology] tools, along with the impact and the associated risks and opportunities" (AACSB 2018, p. 27). In addition, the CPA Evolution, the joint initiative of the National Association of State Boards of Accountancy (NASBA) and the AICPA, will launch in 2024 with changes to the CPA exam and licensure model. The new model is a core-plus-discipline model, where technology concepts are incorporated into all core and discipline sections of the Exam (Behn, 2021). The CPA Evolution reflects the need for today's CPAs to be able to leverage accounting *and* technology skills to solve complex problems early in their careers.

How do accounting professors respond to the challenge of adding new course content without sacrificing core knowledge and skills? One approach is to target courses that currently teach technology *foundation* topics and introduce an *emerging* technology that builds on those topics, with a series of exercises that explicitly connect the foundation knowledge and skills

* Corresponding author at: UNC Wilmington, 601 S. College Road, Wilmington, NC 28403, United States.
E-mail addresses: leel@uncw.edu (L. Lee), casterellag@uncw.edu (G. Casterella).

https://doi.org/10.1016/j.jaccedu.2023.100858
0748-5751/

to the emerging tool. We call this a mental model approach because it encourages students to use their understanding of the foundation topic as the scaffolding for learning the new technology. Leaning on the foundation knowledge de-mystifies the new technology to some degree, while also reinforcing the value of the foundation concepts, since the concepts are revisited through a different user interface.

The purpose of this paper is to demonstrate our mental model approach in the context of an accounting course that has taught relational database concepts and structured query language (SQL) for many years, using Microsoft Access and/or MySQL. We modified the course to add an emerging data analytics tool, Alteryx, in an incremental fashion to illustrate how the two technologies accomplish the same data manipulation tasks but in different ways. The goal is to teach two technologies for data manipulation in a synergistic manner, so that the students' experience in Alteryx reinforces their mental models of SQL queries, and that their understanding of SQL queries will likewise help their understanding of Alteryx.

We chose SQL as the foundation topic because it remains a critical skill for accounting students. Relational databases and SQL have been topics in introductory information systems and AIS textbooks for years (e.g., Rainer & Prince, 2022; Romney et al., 2021), since SQL is the standard language for creating, manipulating, and administering the relational databases that underlie accounting information and enterprise systems. However, SQL's importance extends beyond the operational databases of the organization. SQL is also a topic in data analytic textbooks (e.g., Richardson et al., 2023) because the data manipulation subset of SQL is recognized by public accounting firms and the AICPA as providing a strong technical foundation for data preparation, analytics, and visualization (PwC, 2015; AICPA & NASBA, 2021).

We chose Alteryx Designer as the emerging technology because it is a top-rated data ETL[1] tool for large organizations (Gartner, 2021a) and is also widely adopted among large accounting firms and tax departments (Meyer, 2021; O'Brien & Stone, 2021). Alteryx is an end-to-end analytics platform, with products to support data preparation and analytics, data science and machine learning, process automation, data sharing, and analytic model deployment (Alteryx, 2021). Alteryx Designer is the platform's product for data preparation and analytics. Alteryx Designer provides a no-code, drag-and-drop interface that allows users to create workflows to cleanse, transform, and blend data without writing SQL code.

This paper is organized as follows. In the next section, we explain our instructional approach and our choice of the two technologies to use. Then, we describe our teaching exercises, including the linkages between SQL and Alteryx Designer. Finally, we discuss our implementation and the initial assessment of the effectiveness of this approach.

2. The mental model approach to teaching SQL and Alteryx

2.1. Mental models

Mental models are "cognitive representations of external reality" (Jones et al. 2011, p. 46). They were originally conceptualized by Craik (1943) as the simplified versions of reality we develop in our minds based on our perceptions of how the world works. Johnson-Laird (1983) further developed the concept of mental models to explain how individuals develop these internal representations of the world and use them in their everyday actions and decision-making. Mental models help individuals acquire, store, recall, and decode information about their environment; they also allow individuals to generalize from past experiences as a basis in future decision-making (Jones et al., 2011). Research on mental models has been used in the accounting field in a variety of ways, such as to describe how an auditor forms an overall understanding of a client (Knechel et al., 2010), define what it means to be an accounting professional (Bailey, 1995), and explain how performance measurement systems impact individual performance (Hall, 2011). Educators also recognize the usefulness of mental models as the "scaffolding" on which to build new concepts and skills (van Renaud & Biljon, 2004).

2.2. Structured query language (SQL)

Relational databases have long been a fundamental topic in AIS courses (e.g., Garnsey et al., 2019; Harper & Dunn, 2018; Bain et al., 2002; Romney et al., 2021), particularly since they underly many of the enterprise systems that are the primary source of the organization's financial and managerial accounting data. Relational database topics in traditional AIS courses include tables, primary and foreign keys, normalization, and structured query language (SQL). SQL is the standard programming language for creating, populating, querying, and managing relational databases. Most traditional AIS courses and textbooks focus on querying databases with the SQL SELECT statement, since this statement is used to populate financial and management reports and to create data sets for internal and external auditors and others.

More recently, there has been an increased demand for accounting students with data analytic and emerging technology skills (AACSB, 2018; PwC, 2015; Behn, 2021). While this may sound like a move away from traditional relational database and SQL topics, we find that this is not the case. Data analytics cannot be performed until the source data is extracted from enterprise (and other) systems, prepared, and loaded into a data analytic tool. This extract-transform-load (ETL) process is often performed using SQL or tools that incorporate SQL-like operations. SQL is recognized as a core competency for the extraction and pre-processing of data from enterprise systems (Fotache & Strimbei, 2015) as well as for querying Big Data systems (Silva et al., 2016; Watson, 2014). As Harper and Dunn (2018) emphasize, "Schools must not lose sight of fundamentals ...

[1] ETL stands for Extract-Transform-Load, three key and often time-consuming steps needed prior to data analysis.

data management is an important prerequisite for data analytics: Students who understand sound database design and data models will be better equipped to build dashboards, create visualizations. . .and perform other analytics" (p. 53). The CPA Evolution model curriculum echoes this sentiment with its inclusion of relational database concepts and ETL skills in almost all core and discipline-specific sections (AICPA & NASBA, 2021).

Teaching SQL is important, but also challenging (Vijayasarathy & Casterella, 2016; van Renaud & Biljon, 2004). SQL is a declarative programming language where the query writer must determine which tables contain the relevant source data and visualize how to join rows of data from disparate tables, filter the rows based on multiple criteria, and then aggregate the rows to return the desired dataset (Mitrovic, 1998). These manipulations must be specified within the confines of the SELECT statement with its limited and specifically ordered set of clauses (e.g., SELECT, FROM, WHERE, GROUP BY). This contrasts with other programming languages where the writer specifies the operations to perform on the data, with conventional IF-THEN-ELSE logic as necessary (Matos & Grasser, 2002).

Recognizing the difficulty of teaching SQL, van Renaud and Biljon (2004) used a mental model approach to compare two teaching methods. The first method involved using a query design interface of a specific tool (Microsoft Access). In contrast, the second method used a tool-independent approach where a specific tool was used once the student mastered the basic syntax and semantics of SQL. In this way, the concepts were learned first and then linked to their implementation in a particular tool. van Renaud and Biljon (2004) found that the second method's tool-independent approach focusing on fundamental SQL knowledge before tool exposure emerged as the better approach for teaching SQL.

2.3. Alteryx

Alteryx is a data analytics platform—an integrated set of tools that supports the end-to-end process of data preparation and analytics, with a variety of analytic and machine learning techniques (Alteryx, 2021). The IT research firm Gartner classified Alteryx as the sole "challenger" in the market of data science technology leaders like SAS and IBM (Gartner, 2021b). Many large organizations, including 7-Eleven, Walmart, Coca-Cola, and the Salvation Army, have successfully deployed Alteryx to reduce the hours spent analyzing data and disseminating results.[2] Alteryx is also becoming an important platform for audit and tax work (Meyer, 2021; O'Brien & Stone, 2021), and many of the large accounting firms have partnered with Alteryx to help their clients implement Alteryx-based analytics solutions.[3]

Alteryx Designer is part of the Alteryx platform and is an ETL and data integration product (e.g., Zhang, Porwal, & Eaton, 2020). The drag-and-drop interface allows users to build a visual workflow to cleanse, transform, filter, summarize, and combine source data sets without having to write SQL statements or programming scripts. The workflow also supports predictive analytic and visualization tools. A sample workflow is shown in Fig. 1. The focus of our exercises is on the tools within Alteryx Designer that overlap with SQL queries, i.e., tools in the data input, preparation, and blending areas of the workflow as shown in Fig. 1.

2.4. Model overview

SQL and Alteryx both support dataset filtering, aggregation, and joins, although these similarities may not be apparent to students who are new to SQL and data analytics tools. We follow the mental model approach of van Renaud and Biljon (2004) to guide our instruction:

> *"When teaching any skill we have to take both previous learning and experience into consideration as well as the best possible way of teaching the skill. When learning a new skill, humans form mental models of particular problems and formulate internal patterns for solving these problems. When they encounter the same problem later they apply these previously internalised patterns to solve the problem." (p. 244)*

We use this mental model approach to create an iterative and incremental instructional model for teaching SQL and Alteryx. Our model (Fig. 2) has three layers: the foundation or conceptual layer, the SQL tool layer, and the Alteryx layer.

In the foundation layer, we teach database concepts including data organization with tables that have primary keys and are inter-related via foreign keys. The foundation layer also includes the SQL SELECT statement and how each clause in this statement (e.g., SELECT, FROM, WHERE, GROUP BY) manipulates data from the source table(s) in specific ways.

In the middle layer, students apply the foundation layer concepts using a database management system such as MySQL (an open-source database management system) or Microsoft Access. Students write the SQL code to accomplish specific data manipulation tasks.

In the top layer, we leverage students' understanding of SELECT statements to accelerate their learning of Alteryx Designer for cleansing and manipulating data. The drag-and-drop workflow interface of Alteryx Designer allows students to visualize what the SQL queries accomplish, thus reinforcing the students' mental models of data manipulation while also building agility with respect to emerging technology tools. Specifically, we develop a set of exercises that introduce Alteryx

[2] https://www.alteryx.com/customer-center.
[3] https://www.alteryx.com/alteryx-alliances.

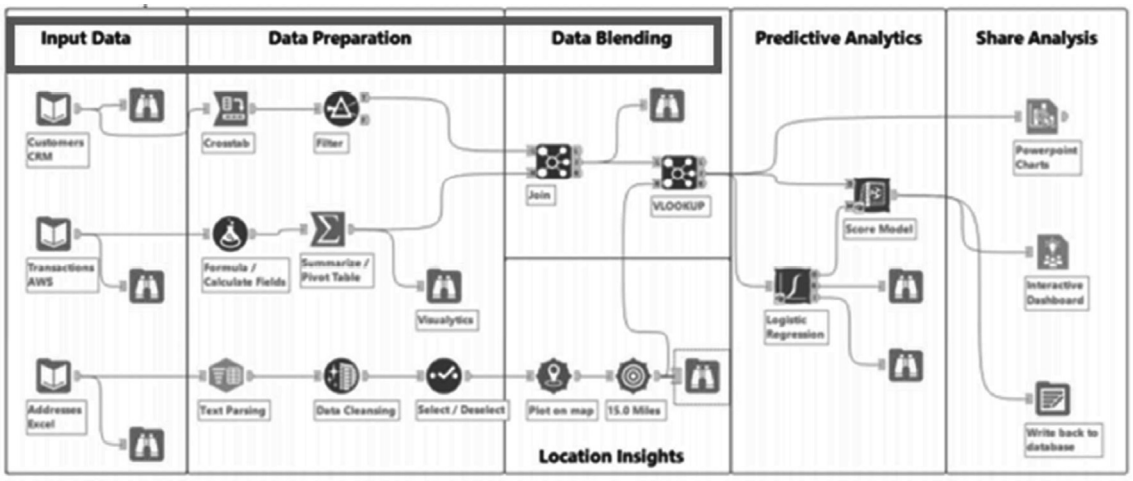

Fig. 1. Components of a data analytics task in Alteryx Designer. *Note:* This figure demonstrates the focus of the exercises in this paper: input data, data preparation, and data blending.. *Source:* https://www.alteryx.com

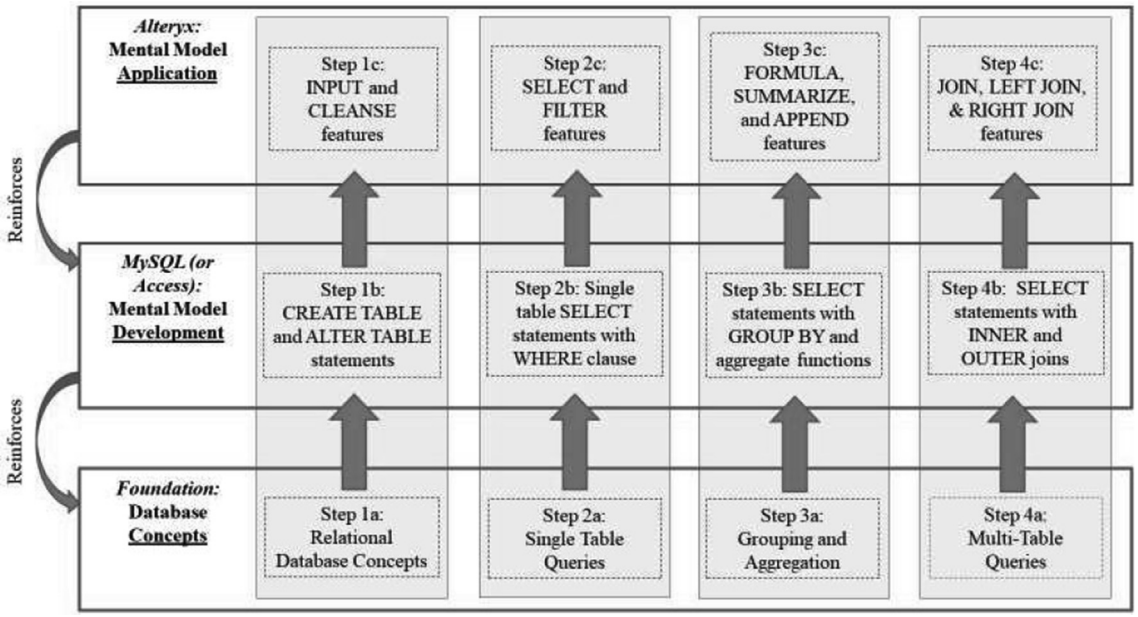

Fig. 2. Iterative, mental model approach to teaching SQL & Alteryx.

Designer from an SQL perspective. With this integration of SQL and Alteryx Designer, the goal is for students to improve their overall competency in data preparation skills for analytics.

As shown in Fig. 2, we iterate through this "Concept-SQL-Alteryx" sequence in four increments or steps that begin with data organization and cover the SQL SELECT statement with the WHERE clause, GROUP BY clause, aggregate functions, and joins. In Step 1a, the instructor introduces traditional database concepts, such as database design, primary and foreign keys, normalization, and 1-to-many relationships. Students apply this learning in Step 1b to create a database using a relational database management system such as MySQL or Access.[4] In Step 1c, students begin to learn Alteryx by inputting data and cleansing data.

[4] In our implementation of this exercise, we used MySQL as the relational database. However, any relational database that supports SQL can be used. The Teaching Notes describe an implementation with MS Access.

Step 2 focuses on data selection and filtering. We explain how a dataset can be manipulated to return only specific columns and specific rows (Step 2a) and illustrate the concepts in SQL by emphasizing the SELECT and WHERE clauses (Step 2b). In Step 2c, the Alteryx tools of "Select" and "Filter" are introduced.

In Step 3, students learn about grouping data together and aggregation functions, such as average, sum, min, and max. In Step 3b, students solve queries using grouping and aggregate functions in SQL. In Step 3c, students use the "Formula," "Summarize," and "Append" tools in Alteryx.

In Step 4, students focus on the concept of joining multiple tables together through inner joins and outer joins. In addition, more advanced joins are discussed such as left joins, right joins, and self-joins. In Step 4b, students join multiple tables with SQL. In Step 4c, students join multiple tables in Alteryx using the "Join" tool and learn how Alteryx implements left and right joins, as well as the "excluding joins" options.[5]

In this iterative process, the students are solving the same problems / queries with both SQL and Alteryx. The students are introduced to the concepts and then work the problems first with SQL to develop their mental model of data manipulation. This is followed by working the same problem / queries with Alteryx so they can apply their mental model to a new environment.

3. The SQL/Alteryx assignment

3.1. Learning objectives

The goal of the exercises is for students to understand how to filter, aggregate, and link data from multiple files/tables to answer specific questions about the data. We want students to understand how to manipulate data in these ways using both SQL SELECT statements and Alteryx workflows. The exercises have the following specific learning objectives:

1) Write SQL queries that include filtering, aggregate functions, grouping, and joining to perform simple data analyses.
2) Create workflows in Alteryx Designer to perform the same analyses as done with SQL.
3) Understand which Alteryx Designer tools relate to the specific clauses or keywords in an SQL query.

3.2. Assignment description

The data used in this assignment is from a simple database for a fictional bookstore (Beach Books). The database contains six tables: book, inventory, location, book-author, author, and publisher. The data for each table is contained in an Excel workbook that can be imported into the relational database.[6] Students must work the questions in both SQL and Alteryx. The questions for each part of the assignment are listed in Table 1.

The Teaching Notes contain an overview of the assignment, including a student-version of the assignment, a grading rubric, and a supplemental tutorial for teaching joins. In addition, supplemental materials available from the authors include: 1) the Excel file with input data for SQL and Alteryx; 2) a PowerPoint presentation with the solution to each of the queries in both SQL and Alteryx; and 3) solutions in Access and Alteryx. These supplemental files can be used by instructors to customize the assignment for their own students and classes and may be obtained by emailing lee-l@uncw.edu.

4. Implementation guidance

4.1. Audience

This assignment can be used at the undergraduate or graduate level in an introductory or advanced accounting information systems (AIS) or accounting analytics class as part of a module on databases and SQL. As such, the exercises assume that the instructor will already be teaching basic SQL and database design as part of the class, with these exercises meant to supplement that instruction. The exercises assume no prior experience with Alteryx Designer.[7] For purposes of the assignment, Alteryx Designer is referred to simply as Alteryx.

[5] In addition to inner joins, Alteryx focuses on left and right "excluding" joins. For example, a left "excluding" join contains records from the L (Left) input that did not join to records from the R (right) input. A full explanation is found in the Teaching Notes.

[6] We intentionally created a small data set for a simple business so that the focus of the assignment is on learning the SQL SELECT statement and the corresponding workflows in Alteryx. There are many other aspects of data analysis that could be incorporated into future versions of the assignment, particularly with larger, more complex data sets, such as extracting data from enterprise systems, checking for errors in the data, and cleansing and standardizing the data prior to analysis.

[7] The instructor had no previous formal training with Alteryx but has taught databases to accounting students for 10+ years. The instructor spent about 4 h learning Alteryx through online resources to obtain a basic level of competency.

Table 1
Assignment questions.

Part 1 Database concepts	Part 2 Single table queries	Part 3 Grouping and aggregate functions	Part 4 Multiple table queries and advanced joins	Part 5 Additional practice (Sales & inventory)
Q1.1: Import the data into a database Q1.2: Import the data into Alteryx	Q2.1: List the book ID and title of each book Q2.2: List the complete publisher table Q2.3: List the name of each publisher with city of "New York City" Q2.4: List the name of each publisher with a city that is NOT "New York City" Q2.5: List the book ID and book title of each book that has a genre of "Dystopian Fiction" Q2.6: List the book ID and book title of each book that has a genre of "Dystopian Fiction" and is a hardcover Q2.7: List the book ID and book title of each book that has a genre of "Dystopian Fiction" OR has a publisher ID of 9 Q2.8: List the book ID, book title, and price of each book with a price between $20 and $30	Q3.1: How many books are of the "Romance" genre? Q3.2: For each book genre, list the genre and the average price Q3.3: What is the title and price of the most expensive book? Q3.4: What is the title and price of the least expensive book? Q3.5: For each type of book (genre), what is the most expensive book? Q3.6: For each type of book, list the type and the average price Q3.7: Compute the discounted price of a book for students (eligible for a 10% discount)	Q4.1: For each book, list the book_ID, title, and publisher name. Order the results by publisher name. Q4.2: For each book published by "HarperCollins Publishers", list the book ID, book title, publisher name, and price Q4.3: List the book title, book ID, and price of each book published by "HarperCollins Publishers" that has a book price of at least $14 Q4.4: List the book ID, book title, and quantity for each book in the "Wrightsville Beach" location Q4.5: List the book title for each book that has the type "Dystopian Fiction" and published by "Katherine Tegen Books". Q4.6: Are there any publishers that don't have an associated book in the book table? Q4.7: Are there any authors in the AUTHOR table that haven't written any books (book_author table)?	Query 5.1: Beach Books' pricing method is to markup hardcover and paperback books 67% over cost, and e-books 40% over cost. Write a query to check the markup for all hardcover and paperback books. The query should list the book title, publisher name, genre, version, price, unit cost, and markup (a calculated field, shown as a %). Query 5.2: What are the top 5 selling hardcover books this year-to-date (YTD), based on quantity sold? List the book title, publisher name, genre, price, unit cost, YTD quantity sold, YTD sales, and YTD cost. Note YTD sales and YTD cost are calculated fields. Query 5.3: Show the YTD total quantity sold and total sales amount for e-books, for hardcover books, and for paperback books. Query 5.4: What is the current value of inventory on hand?

4.2. Implementation hints

The SQL-Alteryx assignments have been implemented in the Spring 1[8] term as part of a graduate-level accounting database class in a southeast regional public university. The class is required for students on a track focused on assurance, analytics, and IT advisory. In our implementation of the assignment, students worked individually; however, these assignments could also be implemented by students working in pairs. Although we used MySQL for the SQL portion of the assignments, any relational database supporting SQL can be used, and our supplementary materials include the solution in Access. Video recordings were created to show students how to navigate the Alteryx interface and to demonstrate the tools applicable for each assignment.[9]

We followed the iterative, incremental process shown in Fig. 2 by first introducing the relational database concepts and then reinforcing the concepts with an SQL exercise and then an Alteryx exercise. This approach was especially helpful in teaching students to join multiple tables. As the interface for Alteryx is visual, students can simply join tables together through clicking and dragging from the appropriate data sources. Students can then compare the visual representation in Alteryx to the corresponding SQL code. This was especially helpful in explaining the concepts of left joins and right joins, as highlighted in the example in Fig. 3.

As part of the supplemental files, a PowerPoint presentation is provided with screenshots of the solution in both Access and Alteryx. This solution file can be modified by the instructor and provided to students as a guide. The instructor can vary the level of guidance provided to students by including parts of the solution in either Access, Alteryx, or both. This is helpful in Access because the SQL can be provided as a screenshot, with students using the screenshot as a guide in developing their own solution. Likewise, the Alteryx solution shows the specific Alteryx tools used in the analysis and can also be used for guidance. This minimizes the amount of direct instructor support required by this assignment in that the screenshots and hints can be provided to students via the PowerPoint.

As noted in Table 2, the Beach Books assignment supplements the topics in a database class targeted at accounting students. Minimal class time is required to prepare the students for each part of the case, assuming that the instructor spends the requisite class time covering the broader database topics.

[8] The assignments have been used in the Spring 1 term of 2021, 2022, and 2023.

[9] These videos were especially important in that the instructor did not have Alteryx installed on the classroom computer. The videos were created to supplement the in-person instruction and demonstrated in more detail how to solve the queries in both SQL and Alteryx. These videos were used primarily in the 2021 term to support students unable to attend in-person due to COVID concerns.

One particular area of difficulty in implementation is the lack of support for Alteryx on Mac systems. Alteryx provides complimentary student and educator licenses. Because we were not able to install Alteryx in the university labs, some students had problems accessing the software in a timely manner. This issue may be addressed in a cloud-based version of Alteryx.[10]

4.3. Grading

The Teaching Notes include a grading rubric with a recommended distribution of points based on both effort (i.e., turning in the Access and the Alteryx file) and correctness. To assess effort, a grader can open the Access / Alteryx files and scan for completeness. To grade the correctness, different options can be used. One option is that the grader can examine each of the queries in the student's Access and Alteryx files. Another option is to create an online quiz in the learning management system (e.g., Canvas) and have the students respond to questions related to each of the problems. A third option is for the students to copy and paste their query results into a Word or PowerPoint[11] file for grading. We have used combinations of these approaches in assessing the assignment. A primary feature of this assignment is that students solve the same problems with both SQL and Alteryx, with the expectation of obtaining the same answer. As such, providing feedback on the correctness of each question is not as critical as it would be with other assignments in that students should already know the correctness of their responses through completion of the assignment.

5. Assignment efficacy

The assignments were first introduced in the Spring 2021 term in a graduate-level AIS class as part of a Master of Science in Accountancy program. All 36 students in the class completed a survey[12] at the end of the term after completing the SQL and Alteryx assignments. We removed one respondent from the analysis since that respondent appeared to have mistakenly selected "strongly disagree" for all the responses instead of "strongly agree," evident from this respondent's open-ended comments which were very positive on the use of Alteryx with no mention of anything negative. With this removal, we have 35 usable responses.

The responses from students (Table 3) were overwhelmingly positive, with 100% of the class responding that they somewhat-to-strongly agreed that Alteryx would be useful in their future careers, their future employers should use Alteryx for data analysis, Alteryx should continue to be part of the course, and that Alteryx was easy to use. Table 3 also summarizes the results of the survey questions related to the specific learning objectives about using both SQL and Alteryx for data analysis. From Table 3, we see that 97% of the respondents somewhat-to-strongly agreed that the Alteryx assignments helped them better understand SQL, the overall relationship among data, how tables are joined together, the different types of join, the selection and filtering of data, the use of aggregate functions in data analysis, and the importance of data cleansing. Over 90% of the respondents somewhat-to-strongly agreed that the Alteryx assignments helped them better understand the connection between the SQL WHERE clause and data filtering.

We next assessed the general usefulness of pairing SQL with Alteryx. About 88% of the respondents said that they somewhat-to-strongly agreed that SQL helped them understand Alteryx (Q14), and 91% of the respondents said that they somewhat-to-strongly agreed that Alteryx helped them understand SQL (Q15). The responses to these two general questions are somewhat lower than the previous, more specific questions, although they still reflect a highly positive perception of the SQL-Alteryx combination. We asked one final question for students to rate their overall perception of the Alteryx exercise from Poor (1) to Excellent (5). All 35 respondents rated the exercises as either "good" (25.71%) or "excellent" (74.3%). Overall, these results show that student perceptions were highly aligned with our learning objectives.

In addition to the survey of student perceptions, we examined student performance on four SQL final exam questions, as shown in Table 4. These same questions were used in two terms, Spring 2020 and Spring 2021. In Spring 2020, Alteryx was not used, and more class time was spent on SQL. In Spring 2021, we implemented the SQL-Alteryx mental model approach. Table 4 shows that student performance was roughly the same for queries requiring inner joins, but performance notably improved on the outer join query for the Spring 2021 (SQL-Alteryx) students. This analysis provides one additional data point that encourages us to refine and improve the mental model approach in future terms, particularly with respect to outer joins. In our experience, outer joins are more challenging for students than inner joins, and Alteryx's visualization tools may allow spending more time on outer join tasks to strengthen students' competence in this area.

The last section of the student survey in Spring 2021 asked for open-ended feedback (positive and negative) on the assignments (Appendix A). A common theme in the positive responses was the ease of use and perceived usefulness of visualization-based analysis enabled by Alteryx, as well as the recognition of the connections between Alteryx and SQL (Table A-1).

[10] A cloud-based version of Alteryx was released in 2022; however, we have not used the cloud-based version with this assignment but would expect it to perform similarly to the desktop version.

[11] As part of the supplemental files, a PowerPoint document is provided with screenshots of the solution in both Access and Alteryx. This file can be modified by the instructor to vary the amount of guidance provided to students for each query. To facilitate grading, the instructor can require students to include screenshots of their own solution in this file.

[12] The survey was reviewed and approved by the IRB at the institution.

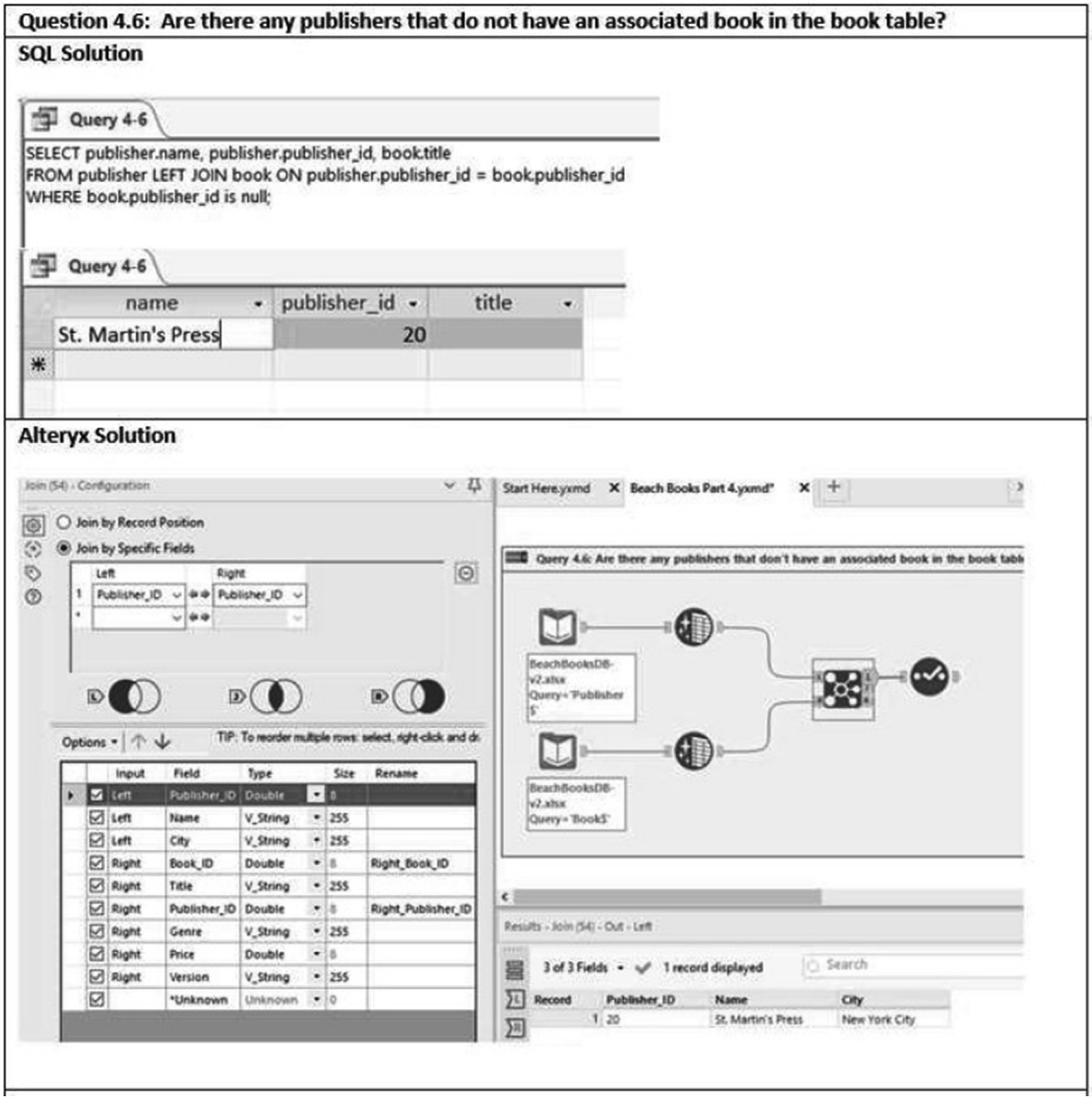

Fig. 3. SQL and Alteryx example with left joins.

Table 2
Recommended class time and approximate assignment completion times.

General database topics[a]	Recommended class time for concepts[b]	Applicable Beach books assignment	Approximate time to complete assignment
Basic database concepts	2–3 Class Periods	Part 1: Importing the data in SQL and Alteryx	30 min to 1 h
Single table queries	1–2 Class Periods	Part 2: Single Table Queries with SQL and Alteryx	~ 1 h
Grouping and aggregate functions	1–2 Class Periods	Part 3: Grouping and Aggregate Functions with SQL and Alteryx	~ 1 h
Multiple table queries and advanced joins	2–3 Class Periods	Part 4: Multiple Table Queries and Advanced Joins with SQL and Alteryx	~ 1 h
Additional practice	Can assign without additional class time	Part 5: Additional Practice with queries related to sales and inventory	~ 1 h

[a] We assume that instructors use outside materials to teach the general database topics and concepts, such as those in AIS / Analytics textbooks (e.g., Romney, Steinbart, Summers, and Wood 2021; Richardson, Terrell, and Teeter 2023).
[b] 75-minute class periods.

Table 3
Student feedback (Multiple choice).

Question	Strongly Disagree	Disagree	Somewhat Disagree	Neither Agree Nor Disagree	Somewhat Agree	Agree	Strongly Agree	Percentage "Somewhat-to-Strongly Agree"	Mean (N = 35)	Std Dev
	(1)	(2)	(3)	(4)	(5)	(6)	(7)			
1. I believe that Alteryx will be useful to me in my future work in audit	0	0	0	0	3	10	22	100.00%	6.543	0.648
2. I believe that Alteryx will be useful to me in my future work in accounting	0	0	0	0	4	10	21	100.00%	6.486	0.692
3. I would recommend that my future employer use Alteryx for data analysis	0	0	0	0	2	12	21	100.00%	6.543	0.602
4. I would recommend that [this class] continue to use Alteryx.	0	0	0	0	0	7	28	100.00%	6.800	0.400
5. I found Alteryx very easy to use.	0	0	0	0	1	5	29	100.00%	6.800	0.466
6. The Alteryx assignments helped me understand SQL better	0	1	0	0	3	16	15	97.14%	6.229	0.959
7. The Alteryx assignments helped me understand the overall relationship among data	0	0	0	1	1	15	18	97.14%	6.429	0.688
8. The Alteryx assignments helped me understand how tables are joined together	0	0	1	0	1	13	20	97.14%	6.457	0.805
9. The Alteryx assignments helped me understand the different types of joins	0	0	0	1	4	13	17	97.14%	6.314	0.785
10. The Alteryx assignments helped me understand the selection and filtering of data	0	0	0	0	0	14	21	100.00%	6.600	0.490
11. The Alteryx assignments helped me understand the connection between the SQL WHERE clause and data filtering	0	1	1	1	3	15	14	91.43%	6.057	1.145
12. The Alteryx assignments helped me understand the use of aggregate functions in data analysis	0	0	0	0	4	19	12	100.00%	6.229	0.636
13. The Alteryx assignments helped me understand the importance of data cleansing	0	0	0	1	2	10	22	97.14%	6.514	0.732
14. SQL helped me better understand Alteryx.	0	2	2	0	7	10	14	88.57%	5.800	1.410
15. Alteryx helped me better understand SQL.	0	2	0	1	6	10	16	91.43%	6.000	1.287
	Poor (1)	Fair (2)	Average (3)	Good (4)	Excellent (5)					
16. Overall, how would you rate the Alteryx exercises?	0	0	0	9	26			100%	4.743	0.437

It is worth noting that many of the students had completed internships and accepted job offers with the Big 4 or regional accounting firms and knew their employers were adopting or using Alteryx. Thus, the perceived relevance and usefulness of

Alteryx were high for our group of graduate students. In addition, at least one of the firms discussed data analytics and Alteryx at a Beta Alpha Psi virtual event in the Spring 2021 term, so that many of our junior and senior undergraduate accounting students also heard about Alteryx. We expect that accounting students at the undergraduate level will have a positive attitude about learning Alteryx, as firms increasingly discuss their adoption of analytics and emerging software tools at various career-related events. At the undergraduate level, we would note that exposure to Alteryx may give students a competitive advantage when they interview for internships.

The only directly negative theme in the students' survey responses was from Mac users or about the technical specifications needed to run Alteryx on laptops.[13] This is a practical consideration that instructors need to address in advance, as the educational license for Alteryx at the time we implemented the class (Spring 2021) was neither Mac-compatible nor cloud-based.

The last open-ended question in the survey was for students to provide additional feedback. Students consistently commented that they enjoyed their learning experience with Alteryx and the associated exercises. A common theme from the students was the ease of learning to manipulate data with Alteryx compared to SQL. Alteryx's drag-and-drop interface was clearly preferred to MySQL's command-line coding or Access' SQL editor or query design grid, although many students saw the tools as reinforcing foundation SQL concepts. The following responses illustrate these points:

- "SQL can be overwhelming to look at and try to understand. Alteryx is more simple to use."
- "Easy compared to SQL."
- "Ease of use. I liked being able to see the process through the workflow."
- "I liked that each step was represented by an icon that helped create a satisfying map at the end."
- "It simplified the joins because you can see them visually (specifically left and right)."
- "I liked how you could see the connections being made, and then getting the answer in SQL was very rewarding."
- "When pairing it [Alteryx] with the MySQL assignments, it was very helpful and interesting to see it in another format."
- "Alteryx helped me understand SQL better. . .especially when we started joining tables."
- "Alteryx was beneficial in understanding the connections between the data and the processes the data was undergoing in our queries. It also allowed for us to check our work."

The ease of learning Alteryx was a positive for the students and is also a positive for instructors wanting to integrate it into a course that currently covers database concepts and SQL. However, the ease of learning could become a double-edged sword, as some students then question why they are learning other technologies that are not as intuitive, such as MySQL or Access. They may also mistakenly view Alteryx as a substitute for a database management system such as MySQL. This issue was reflected in some of the comments, such as:

- "Alteryx was the most up-to-date software we used. A lot of software we use are out-of-date and aren't realistic in what we will be using in the workforce."
- "MySQL is outdated and this [Alteryx] would be of way more benefit."
- "Wish I had more Alteryx instead of Access."
- "I think Alteryx should replace MySQL as the main focus of this class."
- "After learning SQL and Alteryx, I do not like Access as much. I would have loved to do more Alteryx and less Access."
- "More Alteryx and less MySQL."

In addition to student feedback, we shared this assignment with a non-author colleague at another university who teaches a graduate-level accounting analytics course. This instructor has used the Beach Books assignment four times (as of Spring 2023), devoting about two weeks to SQL and the ETL process before assigning the case. When asked for anecdotal feedback regarding the assignment, this instructor provided the following:

> *What I love about this case is that it truly focuses on technological agility and understanding data. Students are able to see how their understanding of data and ETL processes learned during previous course modules (specifically SQL) can be applied in the context of a different application (in this case, Alteryx). With the rapid pace that technology is evolving, I believe it is essential that student learning outcomes involve knowledge and skills and not just tools. I want my students to feel comfortable using any tool. This case shows them that new applications do not have to be scary – that they already possess the knowledge and skills necessary to extract and transform data. After implementing this case, my students are more confident not just with Alteryx but with their abilities to understand and manipulate data. Performance on subsequent assignments using Alteryx improved after using this case as a learning tool.*

The feedback from both students and instructors has been overwhelmingly positive. Although we were pleased with the students' overall response to Alteryx, we will continue to be mindful in future terms to also emphasize the importance of learning the foundation skills, particularly as Alteryx is an expensive product/platform and not every accounting firm and

[13] For the technical requirements to run Alteryx Designer, see https://www.alteryx.com/products/alteryx-platform/alteryx-designer#system-requirements.

Table 4
Comparison of final exam questions (No Alteryx versus Alteryx).

Question type and point value	Question	Spring 2020 (No Alteryx) N = 51		Spring 2021 (Alteryx)N = 36			
		Mean	stdev	mean	stdev	t Stat	P (T<=t) two-tail
Short answer (3 points): Requires joining 4 tables together.	Q1: List the customer number and names (first and last) of all individuals who have purchased books in the 'Fitness' category	2.5098	0.6745	2.6667	0.5345	−1.1271	0.2629
Short answer (3 points): Requires joining 4 tables together.	Q2: What are the titles of the books ordered by Jake Lucas, as well as the order dates?	2.4510	0.6423	2.6389	0.5426	−1.4309	0.1561
Short answer (3 points): Requires joining 3 tables together.	Q3: What is the title of the books written by Anna Adams?	2.4706	0.7029	2.6667	0.5855	−1.3707	0.1741
Short answer (3 points): Requires either a left join or a right join.	Q4: What books in the books table have never been ordered? (Hint: These are books that are in the books table but which are not in the orderitems table)	1.4510	0.8322	2.3056	0.8218	−4.7416	0.0000

certainly not every client will adopt it. The foundation skills (SQL) may be taught using any of several technology tools (e.g., MySQL, Access, SQLite); the key is that the foundation makes it easy to learn new tools that speed up the SQL portion of data preparation and blending.

While both SQL (in Access) and Alteryx may be used for data cleansing, filtering, and summarization tasks, we will also be mindful in future semesters to highlight the differences in scope and purpose between the tools. SQL is a standard implemented in many database management systems, and these database management systems underlie the enterprise systems that support the day-to-day operations of organizations. SQL includes not only a language for manipulating data in these systems, but also for creating the databases, securing them, and administering them. Alteryx, on the other hand, is a data cleansing and analytics tool, which is meant to support decision-making, not daily operations. Our assignment focuses on the tasks that may be supported by either SQL and a relational database or by an analytics tool like Alteryx. However, there are other tasks for which they (i.e., SQL and Alteryx) are not substitutes for each other.

6. Conclusion

Overall, the pairing of SQL with Alteryx is a promising method to incorporate emerging technology into a course that covers relational databases and SQL. Students may feel intimidated and overwhelmed by the new software tools that are mentioned by accounting firms or listed in job postings. One way to build their confidence is to demonstrate that foundational skills do transfer to new technologies, although the connections may not be obvious at first. Alteryx provides a visual workflow approach that has a very different "look and feel" to writing SQL SELECT statements, and yet the Alteryx data preparation tools do parallel SQL query concepts such as filtering, joining, and aggregating. Using SQL as the mental model for learning Alteryx both reinforces the importance of the SQL foundation and de-mystifies new technology tools to make them more approachable.

The incremental mental model approach from Fig. 2 can be extended or adapted in different ways. For example, the SQL-Alteryx exercises could be extended to delve deeper into outer join queries, which are often challenging for students to formulate. More exercises could be added to reinforce the differences between left, right, and full outer joins in SQL and Alteryx. The approach could also be applied to a different set of foundation-plus-emerging technologies. For example, process mining tools are becoming increasingly popular to support audit and compliance tasks (see, for example, the EY Foundation's 2021 process mining curriculum). Process mining tools, such as UiPath or Celonis, are emerging technologies that build on database and process modeling foundations that are often taught in AIS courses. We see potential to use the incremental mental model approach to develop exercises that combining flowcharting or other process modeling topics with Alteryx or process analysis tools such as Celonis. By helping students make explicit connections between foundation concepts and emerging technologies, students can gain confidence in their technology-acquisition skills and see the relevance of the foundation knowledge and skills.

Data availability

Data will be made available on request.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We would like to thank Alexis Sales for her assistance in developing this case and Andrea Kelton for her encouragement and feedback.

Appendix A. Open-ended student comments

	What did you like BEST about SQL / Alteryx assignments?	What did you like LEAST about SQL / Alteryx assignments?
1	*Practicing the use of each icon*	*sometimes the error messages were vague*
2	*The ease of use. I liked being able to see the process through the workflow*	*No complaints. Took up a lot of space and hand to change storage.*
3	*I liked its comparability to excel but its ease of use over excel. Also, I like how you can visually see your data workflow.*	*I initially found it confusing when exporting the data, but once I did it a couple of times I was able to correctly export to the right format.*
4	*They were easy compared to SQL*	*Nothing*
5	*[Alteryx] was like a game- more visually representative compared to MySQL*	*Sometimes the software was slower and it was not compatible with Macs. [The University] should get it on their desktop browser for students to use at home*
6	*Very easy to navigate - user friendly! Much easier than MySQL*	–
7	*It is realistic*	*Some were too long 40 + slides*
8	*I could see why this program is being used in accounting firms*	*Nothing except we should learn more*
9	*Running the data and how easy the data was to understand*	*The formula functions when we have to add new columns*
10	*Easy to use, quick to pick up on*	*I wish we had more Alteryx instead of Access*
11	*That each step was represented by an icon that helped create a satisfying map at the end*	*I would run out of storage and have to figure out how to fix it, which I did.*
12	*Doing MySQL basics and understanding first*	*n/a - I liked it all*
13	*How intuitive the program is and easy to understand*	*n/a. It was really cool*
14	*The ease of the concepts and simple user interface*	*The software seems to have high performance requirements on my laptop*
15	*It simplifying the joins*	*The lack of computers available to us on campus*
16	*-Felt like we were doing something we would use in the future-helpful videos*	*Very difficult for people with Macs to come to campus to use the program, especially with only 2 available computers to use it on*
17	*The system is easy to use out and the functions are self-explanatory. There is little to no guess work with Alteryx.*	*It over simplified the left and right joins. I find them hard to do in MySQL but easy in Alteryx so I would use Alteryx to figure them out and then put them into SQL.*
18	*It was more easier to compare queries in Alteryx as opposed to MySQL*	*Some of the Alteryx Assignments could be time consuming*
19	*How intuitive it was*	*Nothing. I thought they were very useful*
20	*How the workflows were designed for drag and drop, helped visualize what was happening much easier*	*There was nothing in particular I disliked about Alteryx or the assignments*
21	*The ease of clicking and dragging the tools I needed and seeing the SQL code in a visualized manner*	*Overall, I enjoyed them*
22	*I liked how user friendly Alteryx is*	*I didn't really dislike anything about them*
23	*That it was picture based, which helped my learning*	*It took me a while to see the zoom out button so the scrolling around was annoying*
24	*Very easy to use*	*A lot of symbols that were overwhelming at the start*
25	*Visualizing the joins and the helpful videos*	*The very limited access to the program on campus*

Appendix A (*continued*)

	What did you like BEST about SQL / Alteryx assignments?	What did you like LEAST about SQL / Alteryx assignments?
26	*The help slides and videos that explained what each step was doing*	*Some felt a little repetitive at times*
27	*The overall format/style of the Alteryx layout. It's a very intuitive program*	*Nothing. I enjoyed these assignments even more than any other assignments*
28	*The gradual increase in difficulty*	–
29	*I liked how you could see the connections being made to each other and then when getting the answer in SQL was very rewarding*	*Not a whole lot. I wish Alteryx could be used on a Mac, but that's about it. I enjoyed using Alteryx very much.*
30	*I liked how user friendly the Alteryx was.*	*I had no issues. I thought Alteryx was cool*
31	*New experience with a tool being currently used by public accounting firms*	*Some of the assignments seemed redundant like the basic queries*
32	*I liked being able to use a tool that my firm uses regularly*	*I didn't dislike anything about the assignments*
33	*I liked how it was combined with other database programs*	*Firms don't necessarily use Alteryx to just run queries, they use it more to roll-forward workpapers. Implementing additional practical uses of the software would be helpful.*

Appendix B. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jaccedu.2023.100858.

References

AACSB (2018). 2018 Eligibility procedures and accreditation standards for accounting accreditation, https://www.aacsb.edu/-/media/aacsb/docs/accreditation/accounting/standards-and-tables/2018-accounting-standards.ashx?la=en&hash=8DCDA6CE3B0CEF6AB82D39CBF53995DA96111196, accessed 6/8/22.

AICPA & NASBA (2021). CPA evolution model curriculum, https://thiswaytocpa.com/collectedmedia/files/cpa-evolution-model-curriculum-update.pdf, accessed 6/11/2022.

Alteryx (2021). Corporate overview, https://investor.alteryx.com/home/default.aspx, accessed 6/8/22.

Bailey, A. D. Jr, (1995). The practicing professional's mental model: Are we creating the right mental models for new professionals? *Issues in Accounting Education, 10*(1), 191–195.

Bain, C. E., Blankley, A. I., & Smith, L. M. (2002). An examination of topical coverage for the first accounting information systems course. *Journal of Information Systems, 16*(2), 143–164.

Behn, D. (2021). The new CPA evolution CPA exam, *AICPA Insights,* https://www.aicpa.org/news/article/the-new-cpa-evolution-cpa-exam, accessed 6/8/22.

Craik, K. J. W. (1943). *The nature of explanation*. Cambridge, UK: Cambridge University Press.

Fotache, M., & Strimbei, C. (2015). SQL and data analysis. Some implications for data analysis and higher education. *Procedia Economics and Finance, 20*, 243–251.

Garnsey, M., Doganaksoy, N., & Phelan, E. (2019). Topics for the accounting information systems course: A dual perspective approach from educators and employers. *AIS Educator Journal, 14*(1), 36–55.

Gartner (2021a). Data integration tools and reviews, *Gartner Peer Insights*. https://www.gartner.com/reviews/market/data-integration-tools, accessed 6/4/2021.

Gartner (2021b). Magic quadrant for data science and machine learning platforms, https://www.alteryx.com/third-party-content/gartner-magic-quadrant-data-science-machine-learning, accessed 6/4/2021.

Hall, M. (2011). Do comprehensive performance measurement systems help or hinder managers' mental model development? *Management Accounting Research, 22*(2), 68–83.

Harper, C., & Dunn, C. (2018). Building better accounting curricula. *Strategic Finance, 100*(2), 46–54.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, UK: Cambridge University Press.

Jones, N., Ross, H., Lynam, T., Perez, P., & Leitch, A. (2011). Mental models: An interdisciplinary synthesis of theory and methods. *Ecology and Society, 16*(1), 46–59.

Knechel, W., Salterio, S., & Kochetova-Kozloski, N. (2010). The effect of benchmarked performance measures and strategic analysis on auditors' risk assessments and mental models. *Accounting, Organizations and Society, 35*(3), 316–333.

Matos, V., & Grasser, R. (2002). Teaching tip: A simpler (and better) SQL approach to relational division. *Journal of Information Systems Education, 13*(2), 85–88.

Meyer, C. (2021). What faculty should know about Alteryx, *Journal of Accountancy*, 2/9/2021, https://www.journalofaccountancy.com/newsletters/extra-credit/alteryx-data-analysis-for-accounting-faculty.html, accessed 6/4/2021.

Mitrovic, A. (1998). Learning SQL with a computerized tutor. *ACM SIGCSE Bulletin, 30*(1), 307–311.

O'Brien, A., & Stone, D. (2021). A case study in managing the analytics 'iceberg': Data cleaning and management using Alteryx. *Journal of Emerging Technology in Accounting*. https://doi.org/10.2308/JETA-2020-037.

PwC (2015). Data driven – What students need to succeed in a rapidly changing business world, https://www.pwc.com/us/en/faculty-resource/assets/pwc-data-driven-paper-feb2015.pdf, accessed 3/29/2021.

Rainer, R. K., & Prince, B. (2022). *Introduction to Information Systems* (9th edition). NJ: Wiley.

Renaud, K. & van Biljon, J. (2004). Teaching SQL - Which pedagogical horse for this course? *Volume 3112 of Lecture Notes in Computer Science*, 244 - 256. Springer Berlin Heidelberg, Berlin, Heidelberg.

Richardson, V., Terrell, K., & Teeter, R. (2023). *Data analytics for accounting* (3rd edition). New York, NY: McGraw-Hill Education.

Romney, M. B., Steinbart, P. J., Summers, S. L., & Wood, D. A. (2021). *Accounting information systems* (15th edition). Upper Saddle River, NJ: Prentice Hall.

Silva, Y. N., Almeida, I., & Queiroz, M. (2016). SQL: From traditional databases to big data. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, 413-418.

Vijayasarathy, L., & Casterella, G. (2016). The effects of information request language and template usage on query formulation. *Journal of the Association of Information Systems, 17*(10), 674–707.

Watson, H. J. (2014). Tutorial: Big data analytics: Concepts, technologies, and applications". *Communications of the Association for Information Systems, 34*(1), 65.

Zhang, J., Porwal, S., & Eaton, T. (2020). Data preparation for CPAs: Extract, transform, and load," *Journal of Accountancy,* https://www.journalofaccountancy.com/issues/2020/dec/data-preparation-for-accountants.html, accessed 6/8/22.