Journal of Information Systems Education

Volume 33
Issue 2
Spring 2022

Contents and Skills of Data Mining Courses in Analytics Programs

Zinovy Radovilsky and Vishwanath Hegde

Recommended Citation: Radovilsky, Z., & Hegde, V. (2022). Contents and Skills of Data Mining Courses in Analytics Programs. *Journal of Information Systems Education*, 33(2), 182-194.

Article Link: https://jise.org/Volume33/n2/JISE2022v33n2pp182-194.html

Initial Submission: 30 December 2020
Accepted: 21 June 2021
Published: 15 June 2022

ISSN: 2574-3872 (Online) 1055-3096 (Print)

Journal of Information Systems Education, 33(2), 182-194, Spring 2022

Contents and Skills of Data Mining Courses in Analytics Programs

Zinovy Radovilsky
Vishwanath Hegde
Department of Management
California State University, East Bay
Hayward, CA 94542, USA
zinovy.radovilsky@csueastbay.edu, vish.hegde@csueastbay.edu

ABSTRACT

Data Mining (DM) is one of the most offered courses in data analytics education. However, the design and delivery of DM courses present a number of challenges and issues that stem from the DM's interdisciplinary nature and the industry expectations to generate a broader range of skills from the analytics programs. In this research, we identified and compared frequencies of the contents and skills of DM course syllabi in various data analytics programs. We also identified and systemized DM contents and skills in the analytics job market and compared them with the contents and skills from DM syllabi. Based on these analyses and comparisons, we developed four different templates of the DM contents and skills for a DM course at various levels of the analytics education that include: specialized graduate analytics program (MS), general graduate program (MBA), specialized undergraduate analytics program (BS), and general undergraduate program (BSBA). These templates may be specifically useful for educators to design new or improve existing DM courses in data analytics curricula.

Keywords: Business analytics, Data analytics, Data mining, Curriculum design & development

1. INTRODUCTION

For a number of years, data analytics has been demonstrating significant growth and proliferation in industries worldwide (Wymbs, 2016). A 2018 market study of big data analytics by Columbus (2018) shows that a combined big data deployment in enterprises in various industries soared from 17% in 2015 to 59% in 2018. A substantial increase in utilizing data analytics and information technologies corresponded with the need for well-trained professionals to perform data analytics-related tasks. The need of specialists in this area led to a considerable development in the number of academic programs in analytics and, in particular, business data analytics programs. For example, Mills et al. (2016) report a dramatic increase in courses offered in big data analytics (583%) and business data analysis (260%) between 2011 and 2016. Rappa (2019) reports that the number of graduate analytics programs has grown from approximately 30 in 2014 to 110 programs in mid-2019, or more than 3.5 times.

Data analytics is an interdisciplinary field where the management information systems (MIS), decision science, management, and statistics disciplines offer different perspectives on the field. Gorman and Klimberg (2014) define data analytics as an integrated field that combines quantitative methods, information systems and statistics areas. Urbaczewski and Keeling (2019) indicate that the blending of business translator and modeling skills into MIS curriculum has led to the development of analytics programs and concentrations, and many MIS departments have transitioned, at least partially, to analytics. The Certified Analytics Professional (CAP) program launched by INFORMS contains 36 tasks performed by analytics practitioners that fall into seven domains: business problem framing, analytics problem framing, data, methodology selection, model building, deployment, and life-cycle management (Nestler et al., 2012).

Wilder and Ozgur (2015) indicate that the requirements of analytics programs vary across universities; however, many traditional courses from MIS, decision science, management and statistics tend to be a part of these programs. Further, they indicate that a nontraditional component is a course in data mining (DM), which truly differentiates data analytics from the earlier degree programs such as MIS and decision support (DS), thus making data mining a subset of data analytics. Therefore, the design and development of DM courses is an integral part of the development of data analytics programs. As a result, DM has become a subject in the core and/or elective requirements of data analytics programs. According to a web mining study of AACSB-accredited U.S. colleges of business (Zhao & Zhao, 2016), the DM courses are one of the most frequent courses offered by the specialized MS and BS/BA programs in business data analytics (#2 and #3 most frequent, respectively), and #1 most frequent course in MBA concentrations in analytics.

Data Mining is defined as a complex process of examining large sets of data for uncovering hidden patterns and relationships and using them for valuable business insights (Jaggia et al., 2020). The authors also indicate that practitioners generally adopt either CRoss-Industry Standard Process for Data Mining (CRISP-DM) or Sample, Explore, Modify, Model, and Assess (SEMMA) methodologies, which provide the sequential steps involved in DM. The steps include problem identification, data exploration, data collection, data processing, analysis, performance evaluation and implementation. We agree with this DM definition and use it in our research.

The design and delivery of DM in analytics programs pose a number of challenges and issues. First, DM encompasses a broad range of statistical, mathematical, and technical topics, which should be taught to business students with different levels of analytical and technical preparedness. Second, the DM curriculum also needs to ensure that students have the skills to apply DM techniques effectively in the given business context. Therefore, the DM courses should incorporate a broad range of job skills and traits demanded by industry. Third, the DM area is interdisciplinary, thus various business departments or groups of faculty such as information systems, decision sciences, management and others teach DM courses. Finally, DM is taught to students with different levels of business education. For example, Wymbs (2016) discusses the process of infusing data analytics into undergraduate business curriculum, which can be in the form of an analytics concentration in a general business program or a data analytics undergraduate major. As a result, the structure of DM course offerings tends to show high variability depending on the programs they are in and the focus/emphasis faculty chose in terms of analytical topics, technologies, and pedagogy for effective instruction.

Both data analytics and DM education have lately received much attention in the academic research. The papers on DM education focus on the DM topics, technology, data, and pedagogy development. On the other hand, research papers on data analytics emphasize the identification of a broad range of skills required by the analytics job market, but do not go into the details of connecting these skills with the DM course design. In addition, we find wide variability in the definition and list of analytics skills and their prioritization across different levels of analytics programs. Therefore, a natural extension to the existing research in data analytics and DM is to integrate its findings into practical guidelines for the design of DM courses. Our research expands the thought process pursued by Paul and MacDonald (2020) and Radovilsky et al. (2020) in terms of designing DM skills that are aligned with job market requirements and suitable for various levels of data analytics programs. Our research contributes to addressing the practical need that we face as educators in designing and standardizing the DM curriculum in our university, which will be beneficial for DM and data analytics educators at large. Specifically, this research identifies and prioritizes content and skill topics for the DM courses at various business education levels. Our research, therefore, focuses on the Topics part of a sample IS 2010.2 Data and Information Management Core Course of the IS 2010 Curriculum Guidelines (Topi et al., 2010).

The structure of this paper is as follows. We present a review of the literature and objectives of our research in section 2. We analyze in section 3 the DM course syllabi from various levels of data analytics programs and compare skills generated by these syllabi. In section 4, we identify DM skills required by the analytics job market and compare them with the results of DM course syllabi analysis. By integrating the insights obtained in sections 3 and 4, we describe in section 5 the development of contents and skills templates for DM courses at various levels of analytics programs and finish the paper with the conclusions in section 6.

2. LITERATURE REVIEW AND RESEARCH OBJECTIVES

We have organized the review of the existing research on data analytics and DM education into three streams of research papers. The first stream of papers describes the skills generated at different levels of analytics programs by analyzing the description of courses offered in them. Gorman and Klimberg (2014) present curriculum maps of analytics programs, where relative weights attached to courses in undergraduate and graduate degree programs in analytics; the DM subject was present in most of the programs. The authors indicate that the graduate programs were more innovative and attached more emphasis on problem-solving skills, data visualization with a practitioner orientation and emerging analytics topics. Aasheim et al. (2015) conduct comparative analysis of the skills and competencies provided by the undergraduate level data analytics and data science programs. DM courses and related skills were predominant in both types of programs. Wilder and Ozgur (2015) propose a curriculum for undergraduate analytics majors, which includes data management, data visualization, data mining, and descriptive, predictive, and prescriptive analytics. They elaborate on the specifics of the DM course that includes the CRISP-DM process model and DM topics such as classification, clustering, and decision trees. Further, they indicate that the DM course should emphasize framing the business problem, communicating the results in non-technical language, and discussing model implications, impact, and assumptions pertaining to the business problem. Zhao and Zhao (2016) conduct a survey of 215 AACSB accredited business colleges' official websites and report the list of courses offered in the undergraduate/graduate certificate and master's level analytics programs; the DM subject is found to be the most popular across all the programs.

Mills et al. (2016) provide an empirical examination regarding IS programs moving to big data and analytics. They recognize that DM is among the key course offerings in analytics programs. Phelps and Szabat (2017) develop lists of courses and software used, and top ten skills imbedded into the undergraduate data science and data analytics programs; the DM course and related skills appear in both programs. In summary, the discussed research papers characterize analytics programs based on the courses offered in them, and the DM course with the associated skills have emerged as a most common component at all levels of data analytics education.

The second stream of research pertains to the identification of skills sought by the analytics job market and mapping them to data analytics programs. Cegielski and Jones-Farmer (2016) identify a list of skills valued by employers seeking to hire entry-level analytics professionals in a specific job market from undergraduate business programs. They have used the Delphi study of the industry panel, job content analysis and survey of managers to derive this list. The surprising result of their research was that business education, problem-solving skills and business communication had higher frequencies than the data mining quantitative skills. Radovilsky et al. (2018) identify and contrast skills demanded by business data analytics and data science job markets. From the text mining analysis of large pools of related jobs, the authors develop groups (clusters) of skills and classify them into four domains: technical, analytical, business, and communication. They also identify that 5 out of the 10 clusters were in business and communication domains for Business Data Analytics jobs. Pan et al. (2018) indicate that, along with the development of data analytics concentrations, there is an equally important need to improve the data competencies of undergraduate business students who do not major in analytics but still need to have competencies with analytics. They surveyed industry advisors to select the data analytics topics that are appropriate for undergraduate business programs.

Bowers et al. (2018) construct 16 custom topics of analytics job postings and then compare these topics with the skills generated by graduate analytics programs. They found that communication and interpersonal skills (61%), managerial skills (33%), database (33%), and business domain (29%) were the top 4 of the 16 skills constructed from the job postings. In addition, they indicate that less than 5% of the analytics programs generated these skills. Rienzo and Chen (2018) analyze 400 analytical jobs and identify skills demanded by industry that includes SQL, Excel, project management, SAS, database and data mining. The authors also create a list of skills generated by academic programs by surveying 70 academic institutions; the skill list includes analytics core, statistics, data mining, database, project management, and programming.

In addition to revealing the important skills required by the analytics job market, some research and practitioner papers discuss the mismatches (gaps) between the job market requirements and skills students learn in analytics programs and courses. Börner et al. (2018) identify a substantial gap between the soft skills requirements, e.g., communication and writing, in the data science/data engineering job market and the skills generated in the respective analytics programs. Cooper (2021) describes the talent shortages in the data science and analytics job markets that are in part based on not being able to find employees with appropriate skillsets for these jobs. To resolve the gaps between the job market requirements and analytics education, Markow et al. (2017) recommend an accelerated development of new learning pathways which include new data science and analytics degree programs, boot camps, or companies' internal training programs. In summary, the second stream of research papers focuses on identification of a broad range of skills demanded by analytics job market, where business, managerial and communication skills ranked higher. This stream of papers also concentrates on the mismatches between the job market and associated analytics education. However, these research papers do not map analytics job requirements to DM coursework.

The third research stream relates to the design of DM courses. Chakrabarti et al. (2006) present details of a DM curriculum development project that consists of the following design elements: (a) database and data management issues, data process; (b) choice of statistical models, statistical inference considerations and usefulness of the models; (c) algorithm complexity and post processing; and (d) visualization and understanding. Sanati-Mehrizy et al. (2010) describe four DM teaching approaches deployed by undergraduate computer science programs. The mathematical/algorithm and textbook approaches focus on the derivation and intricacies of the various algorithms used in DM with or without a textbook. The topic-based and the applied data mining approaches mix the standard DM topics such as clustering, linear regression, and classification trees with real-world applications. Jafar (2010) explains how he utilized Microsoft Excel's DM add-ins as a front-end to Microsoft's Cloud Computing and SQL Server 2008 business intelligence platforms as a backend to teach a senior-level DM course. Li (2011) describes the application of practice-oriented projects to introduce theoretical DM content to undergraduate business students. He explains how he used the RapidMiner software to assist students in exploring DM processes and algorithms. He also expands the core DM skills with data warehousing, online analytical processing (OLAP) and visual DM.

Wu et al. (2015) illustrate how they taught DM techniques to business undergraduate students using the R software. Asamoah et al. (2017) present an experiential perspective on how a big data analytics course was taught as an elective in the general management and information systems programs. Yap and Drye (2018) describe how they taught the DM tools and techniques. Their approach emphasizes real-world big data sets, data warehouse systems, and DM analytics, which give students the exposure to the world of business where big data is becoming an invaluable asset for strategic decision making, planning, and forecasting. In summary, the third literature stream focuses on selection of the DM topics, applying them using relevant technology and datasets, and the pedagogy for effective delivery to business students. However, these research articles do not incorporate the broad range of skills identified from the job market.

A common theme across the three research streams is the identification of the skills that data analytics and, in particular, DM courses need to generate. However, we find wide variability in the definition and list of skills and their prioritization because these research papers analyze different levels of business education with different purposes/viewpoints. The challenge is how to integrate these research streams and design DM skills for different levels of analytics programs, which is the next logical extension to the body of DM research. We want to discuss two recent articles that advance the research in this direction and try to map analytics job requirements to the analytics curriculum.

Paul and MacDonald (2020) compile a list of job skills and traits identified from the job market by reviewing nineteen research articles, classify the job skills and traits into six groups that are typical in graduate analytics programs, and then map these skills and traits into special designation areas. The DM is one of the special designated areas and the skills and traits that are mapped to this area include association, classification, data munging, insight discoverer, query and analysis, statistical learning, clustering, enterprise analytics, data modeling, time series forecasting, machine learning and several cross-over skills. Radovilsky et al. (2020) indicate that streamlining and standardizing the definition of the DM skills is essential to integrate the findings of the extant research into practical guidelines for designing DM courses. Based on reviewing the pertaining research articles and popular data mining textbooks, they created a standard list of DM skills classified into 22 content-based categories and 8 skill-based categories (see Table 1).

No.	Content-based Categories	Skill-based Categories
1	Dataset Processing, Cleaning and Preparation	Using DM Software
2	Database Management and Data Warehousing	Programming Languages and Coding
3	Data Visualization	Problem Solving
4	Data Reduction/Principal Component Analysis	Apply DM Methods for Business Data
5	Multiple Regression	Selection of Best DM Methods
6	Logistic Regression	Modeling/Decision Making
7	Discriminant Analysis	Interpret and Present Results
8	k-Nearest Neighbors	Written and Oral Communication
9	Decision Trees	
10	Naïve Bayes/Bayes Models	
11	Support Vector Machine	
12	Neural Network	
13	Deep Learning	
14	Cluster Analysis	
15	Association/Affinity Analysis	
16	Collaborative Filtering/Recommender System	
17	Web Data Mining	
18	DM Ethics	
19	Text Mining	
20	Time Series Forecasting	
21	Ensemble Methods	
22	Outlier/Anomaly Detection	

Table 1. DM Course Components

Radovilsky et al. (2020) then analyze frequencies of these content- and skill-based categories across different business data analytics programs and compare them. In addition, they compare the identified frequencies with those demanded by the analytics job market.

In this research paper, we want to integrate these insights to identify DM content- and skill-based categories aligned with analytics job market requirements and suitable for DM courses at different levels of analytics programs. To achieve this goal, we introduce the following three research objectives:

1) Identify and compare frequencies of the content- and skill-based categories taught in DM courses at various data analytics programs.
2) Identify and systemize frequencies of the content- and skill-based categories in business and data analytics job descriptions and compare them with the frequencies found in the first objective.
3) Synthesize the results of the previous two objectives to prioritize contents and skills in DM courses for various data analytics programs.

3. DATA MINING COURSE SYLLABI: ANALYSIS AND COMPARISON

The first objective of our research is to identify and compare frequencies of the content- and skill-based categories taught in DM courses at various levels of analytics programs. We accomplish that by collecting publicly available course syllabi from websites of various business colleges and schools and analyzing their descriptions and contents. Overall, we collected 121 DM course syllabi, which were taught in four academic years from 2016 through 2020. The groups of the sample syllabi are listed in Table 2.

We then analyzed the syllabi for the DM content- and skill-based categories listed in Table 1. Using this data set, we identified and compared frequencies of specific contents and skills in the DM syllabi for: (a) graduate and undergraduate specialized analytics programs; and (b) specialized (graduate and undergraduate) analytics programs and general business programs (MBA and BS/BA) with a concentration in analytics.

Specialized/General	Program Level	Number of Syllabi
Specialized Data Analytics Programs	MS in Analytics	58
	BS in Analytics	31
Analytics Concentrations (General Business Programs)	MBA	18
	Undergraduate (BS/BA)	14
	Total	121

Table 2. Groups of Sample Syllabi

We start our analysis by comparing content- and skill-based categories for the graduate and undergraduate DM syllabi in specialized analytics programs. For that, we identify the frequency of each category in the analyzed graduate and undergraduate course syllabi. In the graduate DM syllabi, these frequencies for the content-based categories vary (in descending order) from the highest of 94.8% for *Cluster Analysis* to the lowest of 5.2% for *Discriminant Analysis*, and for the skill-based categories the frequencies vary from 36.2% for *Modeling/Decision Making* to 17.2% for *Written and Oral Communication*. In the undergraduate DM syllabi, the frequencies for the content-based categories vary from the highest of 96.8% for *Cluster Analysis* to the lowest of 0% for *Web Data Mining* and *Outlier/Anomaly Detection*, and for the skill-based categories from 45.2% for *Using DM Software* to 0% for *Written and Oral Communication*. For each category, we also apply two-tail and one-tail hypothesis testing (significance α value of 0.05) for the difference in category frequencies between the graduate and undergraduate course syllabi. This allows us to identify if the population frequencies of each content- or skill-based category are statistically the same (no significant difference) or if one of the population frequencies is statistically higher than the other for this category.

Based on this analysis, we identified several groups of the content- and skill-based categories that are presented in Table 3. The first group in Table 3 consists of 9 content-based categories (40.9% of the total contents categories), which demonstrate no statistical difference in frequencies in graduate and undergraduate course syllabi. The next group of 6 content-based categories (27.3% of the total contents categories) reveal statistically higher frequencies in graduate vs. undergraduate course syllabi. Most of these categories represent advanced DM contents, e.g., *Neural Network*, *Test Mining*, *Web Data Mining*, and *Ensemble Methods*. Contrary to that, the next group of 4 DM content-based categories (18.2% of the total contents categories), including, among others, *Dataset Processing, Cleaning and Preparation* and *Database Management and Data Warehousing*, contain statistically higher frequencies in the undergraduate courses. This means that the data processing and management contents receive, unconventionally, higher emphasis in the undergraduate analytics programs compared to the graduate ones. The last group of 3 content-based categories (13.6% of the total contents) is comprised of content-based categories with less than 20% of frequencies in both graduate and undergraduate DM courses.

The results in Table 3 demonstrate that 2 out of 8 skill-based categories have no statistically significant difference in frequencies between the graduate and undergraduate course. At the same time, four content-based categories, with the frequencies of 20% or above, are only present in the graduate syllabi, and two more categories contain less than 20% of frequencies on both types of DM course syllabi.

<table>
<tr><th>DM Categories</th><th>Graduate Courses</th><th>Undergraduate Courses</th></tr>
<tr><td colspan="3">Content-based Categories</td></tr>
<tr><td>Cluster Analysis</td><td rowspan="9" colspan="2">No statistically significant difference in frequencies in graduate and undergraduate DM courses</td></tr>
<tr><td>Decision Trees</td></tr>
<tr><td>Association/Affinity Analysis</td></tr>
<tr><td>Multiple Regression</td></tr>
<tr><td>Naïve Bayes/Bayes Models</td></tr>
<tr><td>Data Visualization</td></tr>
<tr><td>Logistic Regression</td></tr>
<tr><td>Collaborative Filtering/Recommender System</td></tr>
<tr><td>Data Reduction/Principal Component Analysis</td></tr>
<tr><td>k-Nearest Neighbors</td><td rowspan="6">Statistically higher frequencies in graduate vs. undergraduate courses</td><td rowspan="6"></td></tr>
<tr><td>Neural Network</td></tr>
<tr><td>Text Mining</td></tr>
<tr><td>Support Vector Machine</td></tr>
<tr><td>Web Data Mining</td></tr>
<tr><td>Ensemble Methods</td></tr>
<tr><td>Dataset Processing, Cleaning and Preparation</td><td rowspan="4"></td><td rowspan="4">Statistically higher frequencies in undergraduate vs. graduate courses</td></tr>
<tr><td>Database Management and Data Warehousing</td></tr>
<tr><td>Time Series Forecasting</td></tr>
<tr><td>DM Ethics</td></tr>
<tr><td>Outlier/Anomaly Detection</td><td rowspan="3" colspan="2">Less than 20% frequencies in both graduate and undergraduate courses</td></tr>
<tr><td>Deep Learning</td></tr>
<tr><td>Discriminant Analysis</td></tr>
<tr><td>Skill-based Categories</td><td></td><td></td></tr>
<tr><td>Using DM Software</td><td rowspan="2" colspan="2">No statistically significant difference in frequencies in graduate and undergraduate courses</td></tr>
<tr><td>Apply DM Methods for Business Data</td></tr>
<tr><td>Modeling/Decision Making</td><td rowspan="4">Present only in graduate courses (frequencies of 20% and up)</td><td rowspan="4"></td></tr>
<tr><td>Interpret and Present Results</td></tr>
<tr><td>Programming Languages and Coding</td></tr>
<tr><td>Problem Solving</td></tr>
<tr><td>Selection of Best DM Methods</td><td rowspan="2" colspan="2">Less than 20% frequency in both graduate and undergraduate courses</td></tr>
<tr><td>Written and Oral Communication</td></tr>
</table>

Table 3. Comparison of Graduate and Undergraduate Course Syllabi

The results in Table 3 demonstrate that 2 out of 8 skill-based categories have no statistically significant difference in frequencies between the graduate and undergraduate course. At the same time, four content-based categories, with the frequencies of 20% or above, are only present in the graduate syllabi, and two more categories contain less than 20% of frequencies on both types of DM course syllabi.

We continue our analysis in this section with a comparison of the content- and skill-based categories and their respective frequencies for DM course syllabi in the specialized analytics programs (MS and BS) and in analytics concentrations of the general business programs, e.g., MBA and BS/BA (Table 4). Similar to the previous results, no statistically significant difference in frequencies was found for 10 out of 22 content-based categories, or 45.5% of the total number of content-based categories (see Table 4).

The results in Table 4 also indicate that the DM syllabi for the specialized analytics courses provide statistically higher frequencies on machine learning and advanced DM content-based categories (5 out of 22 categories, or 22.3% of the total contents), as well as on *Using DM Software* and *Programming Languages and Coding* skill-based categories (2 out of 8 categories, or 25% of the total skills) as opposed to DM syllabi for the general business courses. Contrary to that, only one content-based category, *DM Ethics*, and three skill-based categories, *Apply DM Methods for Business Data*, *Interpret and Present Results*, and *Problem Solving*, are more prevalent (statistically higher frequencies) in the general business courses (see Table 4).

<table>
<tr><th>Data Mining Categories</th><th>Specialized Analytics Courses</th><th>General Business Courses</th></tr>
<tr><td colspan="3">Content-based Categories</td></tr>
<tr><td>Cluster Analysis</td><td colspan="2" rowspan="10">No statistically significant difference in frequencies between specialized analytics and general business courses</td></tr>
<tr><td>Decision Trees</td></tr>
<tr><td>Association/Affinity Analysis</td></tr>
<tr><td>Multiple Regression</td></tr>
<tr><td>k-Nearest Neighbors</td></tr>
<tr><td>Naïve Bayes/Bayes Models</td></tr>
<tr><td>Dataset Processing, Cleaning and Preparation</td></tr>
<tr><td>Data Visualization</td></tr>
<tr><td>Logistic Regression</td></tr>
<tr><td>Collaborative Filtering / Recommender System</td></tr>
<tr><td>Text Mining</td><td rowspan="5">Statistically higher frequencies in specialized analytics vs. general business courses</td><td rowspan="5"></td></tr>
<tr><td>Neural Network</td></tr>
<tr><td>Support Vector Machine</td></tr>
<tr><td>Data Reduction/Principal Component Analysis</td></tr>
<tr><td>Ensemble Methods</td></tr>
<tr><td>DM Ethics</td><td></td><td>Statistically higher frequencies in general business vs. specialized analytics courses</td></tr>
<tr><td>Time Series Forecasting</td><td colspan="2" rowspan="6">Less the 20% frequency in both specialized analytics and general business courses</td></tr>
<tr><td>Database Management and Data Warehousing</td></tr>
<tr><td>Web Data Mining</td></tr>
<tr><td>Deep Learning</td></tr>
<tr><td>Outlier/Anomaly Detection</td></tr>
<tr><td>Discriminant Analysis</td></tr>
<tr><td>Skill-based Categories</td><td></td><td></td></tr>
<tr><td>Modeling/Decision Making</td><td colspan="2">No statistically significant difference in frequencies between specialized analytics and general business courses</td></tr>
<tr><td>Using DM Software</td><td rowspan="2">Statistically higher frequencies in specialized analytics vs. general business courses</td><td rowspan="2"></td></tr>
<tr><td>Programming Languages and Coding</td></tr>
<tr><td>Apply DM Methods for Business Data</td><td rowspan="4"></td><td rowspan="4">Statistically higher frequencies in general business vs. specialized analytics courses</td></tr>
<tr><td>Interpret and Present Results</td></tr>
<tr><td>Problem Solving</td></tr>
<tr><td>Selection of Best DM Methods</td></tr>
<tr><td>Written and Oral Communication</td><td colspan="2">Less the 20% frequency in both specialized analytics and general business courses</td></tr>
</table>

Table 4. Comparison of Specialized Analytics and General Business Courses

4. JOB MARKET CONTENTS AND SKILLS: ANALYSIS AND COMPARISION

The second objective of our research (see section 3) is to identify the content- and skill-based categories demanded by data analytics jobs. We developed a data set of 1,500 unique data analytics job descriptions posted during the 2017-2020 period. We obtained these jobs from various publicly available job-posting websites, including Indeed.com, LinkedIn.com, Monster.com, Dice.com, glassdoor.com, and careerbuilder.com. We also verified the job descriptions manually and included the job records into the data set when the requirements of the position directly related to data analytics. The specified number of jobs signifies a sufficient array of records to be able to perform text mining analytics and obtain statistically significant results from this analysis.

In addition to job titles, we collected data on responsibilities and requirements for each job record. The job responsibilities include a variety of workplace responsibilities and the skills to be required by a prospective employee, e.g., teamwork, oral communication, writing capabilities, reporting, presentation, and problem-solving. The job requirements (or qualifications/skills) include technical skills necessary for that job role, for example, knowledge of specific software like Tableau and Excel, expertise in programming languages like Python and R, and techniques like data visualization and data mining/machine learning methods.

We utilized *R's tm* package with text mining analytics tools (Kwartler, 2017; Shmueli et al., 2019; Weiss et al., 2015) for these 1500 job records. The words extracted from the job descriptions were stored in a structured format referred to as *Text Corpus*, and we cleaned the corpus to isolate common and non-relevant words for our research. Based on the cleaned *Text Corpus*, we developed a *Document Text Matrix (DTM)* that contains actual count and relative frequency of pertinent terms related to DM content- and skill-based categories in the job data set. The 20 most frequent DM content-based terms and corresponding content-based categories from Table 1 are presented in Table 5. The term frequencies in the table vary from the highest of 73.1% for the "Statist" term to the lowest of 10.3% for the "Decis Tree" term.

The results in Table 5 show that 45%, or 9 out of 20 most frequent job-related content terms directly match certain content-based categories in Table 1. This reaffirms the importance of such DM contents such as Database Management, Data Visualization, Multiple Regression, Cluster Analysis, and some others. However, more than one-half of most frequent job-related terms do not have a direct match ("NA" in Table 5) with the content-based categories in Table 1. The explanation of this major difference is that a number of the job-related terms like Statistics/Statistical Analysis, Machine Learning, Large Data, Predictive Analytics, are more general data analytics terms. Therefore, many content-based categories prone to be subsumed into these general analytics terms in Table 5.

The 20 most frequent DM skill-based terms and corresponding skill-based categories from Table 1 are presented in Table 6. The frequencies of these terms in the table vary from the highest of 85.4% for the "report" term to the lowest of 17.1% for the "code write."

No.	Content-based Term		Corresponding Content-based Category from Table 1
	Term	Interpretation	
1	Statist	Statistics/Statistical Analysis	NA
2	machin learn	Machine Learning	NA
3	Database	Database/Database Management	Database Management & Data Warehousing
4	Visual	Visualization/Data Visualization	Data Visualization
5	Larg	Large Data	NA
6	Predict	Predictive Analytics	NA
7	Regress	Regression/Regression Analysis	Multiple Regression
8	Big	Big Data	NA
9	Execut	Execution/Model or Program Execution	NA
10	Cluster	Clustering/Cluster Analysis	Cluster Analysis
11	Mine	Mining/Data Mining	NA
12	Metric	Metrics for Model Performance	NA
13	Queri	Queries/Writing Queries	Database Management
14	trend anal	Trends/Trend Analysis/Forecasting	Time Series Forecasting
15	Evalu	Evaluation/Model Evaluation	NA
16	Deep	Deep Learning	Deep Learning
17	Quantit	Quantitative Models/Analysis	NA
18	Network	Neural Network	Neural Network
19	Intellig	Intelligence/Artificial Intelligence	NA
20	Decis tree	Decision Trees	Decision Trees

Table 5. Twenty Most Frequent DTM Content-based Terms and Corresponding Categories from Table 1

No.	Skill-based Term		Corresponding Skill-based Category from Table 1
	Term	Interpretation	
1	Report	Report/Reporting	Interpret and Present Results
2	Model	Models/Modeling	Modeling/Decision Making
3	Sql	SQL	Programming Languages and Coding
4	Communic	Communication Skills	Written and Oral Communication
5	Excel	Excel	Programming Languages and Coding
6	Python	Python	Programming Languages and Coding
7	Language	Languages/Programming Languages	Programming Languages and Coding
8	Demonstr	Demonstration Skills	Written and Oral Communication
9	Written	Written Communication	Written and Oral Communication
10	Problem	Problem Solving	Problem Solving
11	python r	Python and R	Programming Languages and Coding
12	Tableau	Tableau	Programming Languages and Coding
13	Present	Presentation/Presentation Skills	Interpret and Present Results
14	Implement	Implementation/Model Implementation	NA
15	Verbal	Verbal Communication	Written and Oral Communication
16	Sas	SAS/SAS Applications	Programming Languages and Coding
17	Decis	Decision(s)/Decision Making	Problem Solving
18	Hadoop	Hadoop	NA
19	Spark	Spark	NA
20	code write	Code Writing/Coding	Programming Languages and Coding

Table 6. Twenty Most Frequent DTM Skill-based Terms and Corresponding Categories from Table 1

Most of the skill-based terms in Table 6 (17 out of 20 terms) can be directly matched with the skill-based categories in Table 1. The importance of the *Programming Languages and Coding skill-based* category is evident from its association with 8 job-related terms like *SQL*, *Excel*, *Python*, *Languages/Programming Languages*, and some others. Similarly, the *Written and Oral Communication* skill-based category is directly related to several skill-based terms in Table 6. As can be seen from Table 6, *Communication Skills* is the fourth most frequent term among the skill-based terms. According to Radovilsky et al. (2020), *Communication Skills* term was used in 71.8% of job requirements in data analytics jobs.

Our second research objective also incorporates a comparison of the two sets of results, i.e., the content- and skill-based categories of DM course syllabi and those of the job market requirements (Table 7). We found that many of the content-based categories from DM syllabi analyzed in section 3 were not identified from the job requirements in data analytics (see Table 5), and thus could not be compared in Table 7.

The compared frequencies in Table 7 produced several important observations. First, the DM syllabi have statistically higher frequencies in 4 content-based categories as compared to the job market requirements; the only exception is the *Database Management & Data Warehousing* content-based category, which holds a higher frequency in the job market. Second, the job requirements contain noticeably higher frequencies for a number of skill-based categories compared to those for the DM syllabi. In other words, the DM courses may underemphasize the skill-based categories like *Modeling/Decision Making*, *Using DM Software*, *Programming Languages and Coding*, and *Written and Oral Communication*.

5. DESIGN OF CONTENTS AND SKILLS IN DM COURSES

The third and final objective of our research (see section 2) is to design the contents and skills for DM courses in various data analytics programs. We integrate the results in sections 3 and 4 to design templates of contents and skills for DM courses. The general principles of designing a template are as follows:

1) The content and skill-based categories in a DM course are directly associated with the standard list of these categories presented in Table 1 (see section 2).
2) The design should incorporate the differences in contents and skills of DM courses in various data analytics programs, which we analyzed in section 3.
3) The design should incorporate the frequent content- and skill-based terms identified from the job market requirements (see section 4). In particular, the skill-based terms underrepresented in DM syllabi must be an integral part of this design.

In each designed template, we allocate the content- and skill-based categories into three main groups: *Core*, *Program-specific*, and *Additional* categories. The first two groups are required for each DM course, and the third one is optional. The *Core* categories reflect the most common contents and skills for a DM course in various analytics and general business programs. They would typically include the high-frequency content- and skill-based categories with no statistically significant differences between DM courses in various analytics

DM Categories	Frequency in DM Syllabi	Frequency in Job Market Requirements
Content-based Categories		
Data Visualization	No statistically significant difference in frequencies in DM syllabi vs. job market requirements	
Time Series Forecasting		
Deep Learning		
Cluster Analysis	Statistically higher frequencies in DM syllabi vs. job market requirements	
Decision Trees		
Multiple Regression		
Neural Network		
Database Management & Data Warehousing		Statistically higher frequencies in job market requirements vs. DM syllabi
Skill-based Categories		
Interpret and Present Results	No statistically significant difference in frequencies in DM syllabi vs. job market requirements	
Problem Solving		
Modeling/Decision Making		Statistically higher frequencies in job market requirements vs. DM syllabi
Using DM Software		
Programming Languages and Coding		
Written and Oral Communication		

Table 7. DM Categories in Syllabi and Job Market Requirements

programs, e.g., specialized analytics vs. general business courses (see Tables 3, 4, and 7 for these categories). The *Program-specific* categories are applied in DM courses for a particular type of data analytics program in which these content- and skill-based categories have statistically higher frequencies than in other programs (see Tables 3 and 4 for these categories). The *Additional* categories are optional DM contents and skills that are not allocated in the first two categories and typically have less than 20% frequencies in DM courses in the specific analytics programs. They can be flexibly added to a DM course depending on an instructor's preference and a department from which this DM course is offered.

In Table 8, we present two templates of content- and skill-based categories for graduate DM courses. One of the templates represents a graduate course in a specialized M.S. program in analytics, and the other one is for a graduate DM course in an analytics concentration of a general business program like MBA. As can be seen from Table 8, the *Core* categories in both courses are mostly the same high-frequency content- and skill-based categories, for which there is no statistically significant difference in frequencies between specialized analytics and general business DM courses. In addition, even though our analysis indicated that *Written and Oral Communication* skill is present in less than 20% of DM syllabi, we elevated this skill to the *Core* category because of its importance in the job market requirements discussed in section 4 (see Table 6).

There are, however, significant differences in *Program-specific categories* in the two course templates (see Table 8). In particular, the DM course in specialized analytics programs includes those content- and skill-based categories that have statistically higher frequencies in these programs vs. general business programs. The DM course contents for a specialized analytics program is more intense with a focus on advanced DM methods like *Text Mining*, *Neural Networks*, *Support Vector Machine*, and *Ensemble Methods*, whereas the DM course in a general business program contains only one content-based category of *DM Ethics*. The *Program-specific* skills for the DM course in a specialized analytics program include advanced computer skills like *Using DM Software* and *Programming Languages and Coding*. The latter skill was also recognized as one of the most frequent skills in job market requirements (see Table 6). Conversely, the skills in the DM course for a general business program encompass more business-related skills like *Apply DM Methods for Business Data*, *Interpret and Present Results*, *Problem Solving*, and *Selection of Best DM Methods*.

Various contents and skills are allocated to the *Additional categories* in the described course templates (see Table 8). As previously mentioned, these are DM contents and skills that are not assigned to the required *Core* and *Program-specific* categories in the respective templates and may have less than 20% frequencies in the DM courses in specialized analytics or general business programs. For example, the additional contents in the DM course for a general business program include *Text Mining*, *Neural Network*, *Support Vector Machine*, and *Ensemble Methods*, which are not listed for the DM course for a specialized analytics program. At the same time, skills like *Apply DM Methods for Business Data*, *Interpret and Present Results*, *Problem Solving*, and *Selection of Best DM Methods* are used as additional skills in the DM course for a specialized analytics program. Contrary to that, the DM skills in a general business program emphasize *Using DM Software* and *Programming Languages and Coding* in the *Additional* category of the DM templates.

The template of DM content- and skill-based categories for undergraduate specialized and general business courses are listed in Table 9. As previously discussed for the graduate course templates, the *Core* categories for the undergraduate templates are also the same. However, the *Core* contents have been expanded by adding *Dataset Processing, Cleaning and Preparation*, *Database Management and Data Warehousing*, and *Time Series Forecasting* contents, which have statistically higher frequencies in undergraduate vs. graduate courses. The structure for contents and skills in the *Program-specific* and *Additional* categories of the undergraduate course templates remain similar to those described for the graduate course templates. For example, the *Program-specific* category for the

Category	Graduate DM Course in Specialized Analytics Program		Graduate DM Course in General Business Program	
	Content-based Categories	Skill-based Categories	Content-based Categories	Skill-based Categories
Core	Introduction and Data Mining Process		Introduction and Data Mining Process	
	Cluster Analysis	Modeling/Decision Making	Cluster Analysis	Modeling/Decision Making
	Decision Trees	Written and Oral Communication	Decision Trees	Written and Oral Communication
	Association/Affinity Analysis		Association/Affinity Analysis	
	Multiple Regression		Multiple Regression	
	k-Nearest Neighbors		k-Nearest Neighbors	
	Naïve Bayes/Bayes Models		Naïve Bayes/Bayes Models	
	Dataset Processing, Cleaning and Preparation		Dataset Processing, Cleaning and Preparation	
	Data Visualization		Data Visualization	
	Logistic Regression		Logistic Regression	
	Collaborative Filtering/ Recommender System		Collaborative Filtering/ Recommender System	
Program-specific	Text Mining	Using DM Software	DM Ethics	Apply DM Methods for Business Data
	Neural Network	Programming Languages and Coding		Interpret and Present Results
	Support Vector Machine			Problem Solving
	Data Reduction/Principal Component Analysis			Selection of Best DM Methods
	Ensemble Methods			
Additional	DM Ethics	Apply DM Methods for Business Data	Text Mining	Using DM Software
	Time Series Forecasting	Interpret and Present Results	Neural Network	Programming Languages and Coding
	Database Management and Data Warehousing	Problem Solving	Support Vector Machine	
	Web Data Mining	Selection of Best DM Methods	Data Reduction/Principal Component Analysis	
	Deep Learning		Ensemble Methods	
	Outlier/Anomaly Detection		Time Series Forecasting	
	Discriminant Analysis		Database Management and Data Warehousing	
			Web Data Mining	
			Deep Learning	
			Outlier/Anomaly Detection	
			Discriminant Analysis	

Table 8. DM Specialized and General Graduate Courses

DM course contents in a specialized undergraduate analytics program covers advanced DM methods like *Text Mining*, *Neural Networks*, *Support Vector Machine*, and *Ensemble Methods*, whereas the DM course in a general undergraduate business program contains only one content-based category of *DM Ethics*. The *Program-specific* skills for the DM course in a specialized undergraduate analytics program include advanced computer skills like *Using DM Software* and *Programming Languages and Coding*. Conversely, the skills in the DM course for a general undergraduate business program encompass more business-related skills like *Apply DM Methods for Business Data*, *Interpret and Present Results*, *Problem Solving*, and *Selection of Best DM Methods*.

	Undergraduate DM Course in Specialized Analytics Program		Undergraduate DM Course in General Business Program	
Category	Content-based Categories	Skill-based Categories	Content-based Categories	Skill-based Categories
Core	Introduction and Data Mining Process		Introduction and Data Mining Process	
	Cluster Analysis	Modeling/Decision Making	Cluster Analysis	Modeling/Decision Making
	Decision Trees	Selection of Best DM Methods	Decision Trees	Selection of Best DM Methods
	Association/Affinity Analysis	Written and Oral Communication	Association/Affinity Analysis	Written and Oral Communication
	Multiple Regression		Multiple Regression	
	Naïve Bayes/Bayes Models		Naïve Bayes/Bayes Models	
	Data Visualization		Data Visualization	
	Logistic Regression		Logistic Regression	
	Collaborative Filtering /Recommender System		Collaborative Filtering /Recommender System	
	Dataset Processing, Cleaning and Preparation		Dataset Processing, Cleaning and Preparation	
	Database Management and Data Warehousing		Database Management and Data Warehousing	
	Time Series Forecasting		Time Series Forecasting	
Program-specific	Text Mining	Using DM Software	DM Ethics	Apply DM Methods for Business Data
	Neural Network	Programming Languages and Coding		Interpret and Present Results
	Support Vector Machine			Problem Solving
	Data Reduction/Principal Component Analysis			
	Ensemble Methods			
Additional	k-Nearest Neighbors	Apply DM Methods for Business Data	k-Nearest Neighbors	Using DM Software
	DM Ethics	Interpret and Present Results	Text Mining	Programming Languages and Coding
	Web Data Mining	Problem Solving	Neural Network	
	Deep Learning		Support Vector Machine	
	Outlier/Anomaly Detection		Data Reduction/Principal Component Analysis	
	Discriminant Analysis		Ensemble Methods	
			Web Data Mining	
			Deep Learning	
			Outlier/Anomaly Detection	
			Discriminant Analysis	

Table 9. DM Specialized and General Undergraduate Courses

6. CONCLUSIONS

Data mining is an important non-traditional component that differentiates data analytics from traditional business degree programs and concentrations. As a result, the DM course is one of the most offered courses in various analytics programs. Design and delivery of DM courses in business education pose many challenges and issues that stem from the DM's interdisciplinary nature loaded with technical contents and skills, and also from the evolving industry expectations to generate a broader range of skills from the analytics programs. The goal of this research was to design templates of contents and skills for DM courses that are aligned with the job market requirements and suitable for various levels of data analytics programs.

We accomplish this goal by fulfilling the three research objectives in the following ways. First, we identified and compared frequencies of the content- and skill-based categories of DM course syllabi in various analytics programs. Second, we provided unique research results for identifying and systemizing DM contents and skills in job market requirements for data analytics, and then compared them with the content- and skill-based categories in the DM syllabi. Finally, we developed templates of content- and skill-based categories for the DM courses at various analytics programs.

The outcomes of this research contribute to the theory and practice of data analytics education and, specifically, the DM curriculum, which is a crucial component of analytics programs in business. We designed concrete course templates for DM contents and skills through triangulation of different streams of research and incorporating results of our analysis and comparison of existing DM syllabi and job market requirements. In this process, we applied the standard list of content- and skill-based categories that can channel the requirements of the job market to data analytics curriculum and then to DM course design. From the practitioner's viewpoint, we presented the tangible DM contents and skills templates that can be very useful to design new and improve existing DM courses in various analytics programs.

Despite the described contributions of this research, it also contains several limitations. First, our research and findings are based on a convenience sample of publicly available syllabi. Subsequent research with more sample syllabi is required to further assert our findings and generalizing the results. Second, the DM education is an emerging area, and we anticipate changes as we move forward with this education. There may be a need for a longitudinal study to understand the changes of the DM job requirements and DM syllabi over time and to adjust contents and skills in the DM courses accordingly.

This research can be extended in the future in several ways. We may analyze the DM content- and skill-based categories in conjunction with several important job-related characteristics and attributes, i.e., years of experience, education, job location, and some others. In addition, we can also analyze and compare DM teaching in various analytics programs utilizing the proportion of time and depth of teaching specific content- and skill-based categories and utilize the results of these analyses for further design of DM courses.

7. REFERENCES

Aasheim, C. L., Williams, S., Rutner, P., & Gardiner, A. (2015). Data Analytics vs. Data Science: A Study of Similarities and Differences in Undergraduate Programs Based on Course Descriptions. *Journal of Information Systems Education*, 26(2), 103-116.

Asamoah, D. A., Sharda, R., Zadeh, A. H., & Kalgorta, P. (2017). Preparing a Data Scientist: A Pedagogic Experience in Designing a Big Data Analytics Course. *Decision Sciences Journal of Innovative Education*, 15(2), 161-190.

Börner, K., Scrivner, O., Gallant, M., Ma, S., Liu, X., Chewning, K., Wu, L., & Evans, J.A. (2018). Skill Discrepancies between Research, Education, and Jobs Reveal the Critical Need to Supply Soft Skills for The Data Economy. *Proceedings of the National Academy of Sciences*, 115(50), 12630-12637.

Bowers, M. R., Camm, J. D., & Chakraborty, G. (2018). The Evolution of Analytics and Implications for Industry and Academic Programs. *Interfaces*, 48(6), 487-499.

Cegielski, C. G., & Jones-Farmer, L. A. (2016). Knowledge, Skills, and Abilities for Entry-Level Business Analytics Positions: A Multi-Method Study. *Decision Sciences Journal of Innovative Education*, 14(1), 91-118.

Chakrabarti, S., Ester, M., Fayyad, U., Gehrke, J., Han, J., Morishita, S., Piatetsky-Shapiro, G., & Wang, W. (2006). Data Mining Curriculum: A Proposal (Version 1.0). *The Association for Computing Machinery's Special Interest Group on Knowledge Discovery and Data Mining*. https://www.kdd.org.

Columbus, L. (2018). Big Data Analytics Adoption Soared in The Enterprise In 2018. *Forbes*. https://www.forbes.com/sites/louiscolumbus/2018/12/23/big-data-analytics-adoption-soared-in-the-enterprise-in-2018/#57e9a728332f.

Cooper, K. (2021). How Can We Fix the Data Science Talent Shortage? https://www.springboard.com/blog/data-science/data-science-talent-shortage.

Gorman, M. F., & Klimberg, R. K. (2014). Benchmarking Academic Programs in Business Analytics. *Interfaces,* 44(3), 329-341.

Jafar, M. J. (2010). A Tools-based Approach to Teaching Data Mining Methods. *Journal of Information Technology Education: Innovations in Practice*, 9(1), 1-24.

Jaggia, S., Kelly, A., Lertwachara, K., & Chen, L. (2020). *Business Analytics: Communicating with Numbers* (1st Ed). McGraw-Hill Education.

Kwartler, T. (2017). *Text Mining in Practice with R*. New York, NY: John Wiley & Sons.

Li, D. (2011). A Practice-Oriented Approach to Teaching Undergraduate Data Mining Course. *American Society for Engineering Education*, 1-11. https://peer.asee.org/a-practice-oriented-approach-to-teaching-undergraduate-data-mining-course.pdf

Markow W., Braganza, S., Taska, B., Miller, S. M., & Hughes, D. (2017). The Quant Crunch: How the Demand for Data Science Skills is Disrupting the Job Market. https://www.ibm.com/downloads/cas/3RL3VXGA.

Mills, R. J., Chudoba, K. M., & Olsen, D. H. (2016). IS Programs Responding to Industry Demands for Data Scientists: A Comparison Between 2011-2016. *Journal of Information Systems Education*, 27(2), 131-140.

Nestler, S., Levis, J., & Klimack, B. (2012). Certified Analytics Professional. *ORMS-Today*, 39(5). https://www.informs.org/ORMS-Today/Public-Articles/October-Volume-39-Number-5/Certified-Analytics-Professional.

Pan, K., Blankley, A. I., Mazzei, M. J., Lohrke, C. F., Marshall, J. B., & Carson, C. M. (2018). Surveying Industry Advisors to Select Data Analytics Topics for All Business Majors. *The International Journal of Management Education,* 16(3), 483-492.

Paul J. A., & L. MacDonald. (2020). Analytics Curriculum for Undergraduate and Graduate Students. *Decision Sciences Journal of Innovative Education*, 18(1), 22-58.

Phelps, A. L., & Szabat, K. A. (2017). The Current Landscape of Teaching Analytics to Business Students at Institutions of Higher Education: Who Is Teaching What? *The American Statistician*, 71(2), 155-161.

Radovilsky, Z., Hegde, V., Acharya, A., & Uma, U. (2018). Skills Requirements of Business Data Analytics and Data Science Jobs: A Comparative Analysis. *Journal of Supply Chain and Operations Management*, 16(1), 82-101.

Radovilsky, Z., Hegde, V., & Damle, L. H. (2020). Data Mining in Business Education: Exploratory Analysis of Course Data and Job Market Requirements. *Journal of Supply Chain and Operations Management*, 18(1), 119-139.

Rappa M. (2019). Graduate Degree Programs in Analytics and Data Science. https://analytics.ncsu.edu/?page_id=4184.

Rienzo, T. & Chen, K. (2018). Planning for Low End Analytics Disruptions in Business School Curricula. *Decision Sciences Journal of Innovative Education*, 16(1), 23-41.

Sanati-Mehrizy, R., Sanati-Mehrizy, P., Minaie, A., & Dean, C. (2010). How a Data Mining Course Should Be Taught in an Undergraduate Computer Science Curriculum. *American Society for Engineering Education*, 1-7. https://peer.asee.org/how-a-data-mining-course-should-be-taught-in-an-undergraduate-computer-science-curriculum.pdf

Shmueli, G., Bruce, P. C., Yahav, I., Patel, N. R., & Lichtendahl, K. C. (2019). *Data Mining for Business Analytics: Concepts, Technologies, and Applications in R*. New York, NY: John Wiley & Sons.

Topi, H., Valacich, J. S., Wright, R. T., Kaiser, K., Nunamaker, Jr., J. F., Sipior, J. C., & Jan de Vreede, G. (2010). IS 2010: Curriculum Guidelines for Undergraduate Degree Programs in Information Systems. *Communications of the Association for Information Systems*, 26(18), 359-428.

Urbaczewski, A., & Keeling, K. B. (2019). The Transition from MIS Departments to Analytics Departments. *Journal of Information Systems Education*, 30(4), 303-310.

Weiss, S.M., Indurkhya, N., & Zhang, T. (2015). *Fundamentals of Predictive Text Mining* (2nd ed.). London: Springer-Verlag.

Wilder, C. R., & Ozgur, C. O. (2015) Business Analytics Curriculum for Undergraduate Majors. *INFORMS Transactions on Education,* 15(2), 180-187.

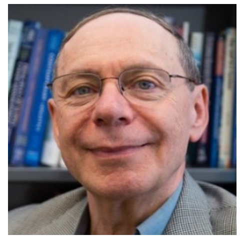
Wu, C., Mai, F., & Yu, Y. (2015). Teaching Data Mining to Business Undergraduate Students Using R. *Business Education Innovation Journal*, 7(2), 64-73.

Wymbs, C. (2016). Managing the Innovation Process: Infusing Data Analytics into the Undergraduate Business Curriculum (Lessons Learned and Next Steps). *Journal of Information Systems Education*, 27(1), 61-74.

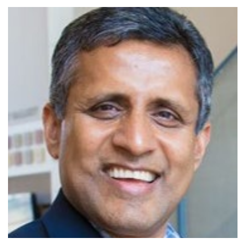
Yap, Y. Y., & Drye, S. L. (2018). The Challenges of Teaching Business Analytics: Finding Real Big Data for Business Students. *Information Systems Education Journal*, 16(2), 41-50.

Zhao, J., & Zhao, S. (2016). Business Analytics Programs Offered by AACSB-accredited U.S. Colleges of Business: A Web Mining Study. *Journal of Education in Business*, 91(6), 327-337.

AUTHOR BIOGRAPHIES



Zinovy Radovilsky is a professor of management at the College of Business and Economics (CBE), California State University, East Bay. He is currently serving as the coordinator of Operations Management and Analytics in the Department of Management. He played a pivotal role in developing and launching a successful Master of Science in Business Analytics program in CBE. Dr. Radovilsky's research interests include business data analytics, time series forecasting, operations and supply chain optimization, education and curriculum development in business analytics and supply chain management. He has published 37 research papers in peer-reviewed journals, four book chapters, and one textbook.



Vishwanath Hegde is a professor of management at the College of Business and Economics, California State University, East Bay. He has over 30 years of experience in a variety of roles and domains that spans both academia and industry. He earned his Ph.D. in Operations Management from the University of Pittsburgh and then worked in IT inductry including ERP software company PeopleSoft. He has been conducting empirical research and has published 20 refereed research articles. His research interests include project management, supply chain finance, online education and business analytics and supply chain management curriculum. He is currently serving as the Interim Associate Dean of Graduate Programs and Research.

Information Systems & Computing Academic Professionals

Education Special Interest Group

STATEMENT OF PEER REVIEW INTEGRITY

All papers published in the ***Journal of Information Systems Education*** have undergone rigorous peer review. This includes an initial editor screening and double-blind refereeing by three or more expert referees.

ISSN 2574-3872