INFORMATION SYSTEMS EDUCATION JOURNAL

Volume 22, No. 2
May 2024
ISSN: 1545-679X

In this issue:

iscap
Information Systems & Computing Academic Professionals

The **Information Systems Education Journal** (ISEDJ) is a double-blind peer-reviewed academic journal published by **ISCAP** (Information Systems and Computing Academic Professionals). Publishing frequency is five times per year. The first year of publication was 2003.

ISEDJ is published online (https://isedj.org). Our sister publication, the Proceedings of the ISCAP Conference (https://iscap.us/proceedings) features all papers, abstracts, panels, workshops, and presentations from the conference.

The journal acceptance review process involves a minimum of three double-blind peer reviews, where both the reviewer is not aware of the identities of the authors and the authors are not aware of the identities of the reviewers. The initial reviews happen before the ISCAP conference. All papers, whether award-winners or not, are invited to resubmit for journal consideration after applying feedback from the Conference presentation. Award winning papers are assured of a publication slot; however, all re-submitted papers including award winners are subjected to a second round of three blind peer reviews to improve quality and make final accept/reject decisions. Those papers that are deemed of sufficient quality are accepted for publication in the ISEDJ journal. Currently the target acceptance rate for the journal is under 35%.

Information Systems Education Journal is pleased to be listed in the Cabell's Directory of Publishing Opportunities in Educational Technology and Library Science, in both the electronic and printed editions. Questions should be addressed to the editor at editor@isedj.org or the publisher at publisher@isedj.org. Special thanks to volunteer members of ISCAP who perform the editorial and review processes for ISEDJ.

INFORMATION SYSTEMS EDUCATION JOURNAL

Editors

Analytics for an Audience of Healthcare Professionals: Curriculum Design and Student Perceptions

Jennifer Xu
jxu@bentley.edu

Monica Garfield
mgarfield@bentley.edu

Department of Computer Information Systems
Bentley University
Waltham, MA 02452, USA

Abstract

There has been an increasing demand for healthcare analytics skills and competence by healthcare organizations. Although many universities have established programs and courses on healthcare analytics, most of these curricula have been designed for information systems (IS), information technology (IT), or analytics students. It is unclear how these curricula would fit the needs of healthcare professionals who have little IT knowledge and background yet also need analytics for their clinical or administrative job roles. This research reports on the design of an executive MBA course intended for an audience of healthcare professionals. The learning objectives, topic coverage, software tools, and assessment methods are presented along with students' perceptions of these aspects of the course. Several important lessons learned are shared and future directions are proposed, which can help other educators design similar healthcare analytics courses for professional audiences.

Keywords: Curriculum design, healthcare analytics, healthcare professionals, student perceptions, data visualization.

Recommended Citation: Xu, J., Garfield, M., (2024). Analytics for an Audience of Healthcare Professionals: Curriculum Design and Student Perceptions. *Information Systems Education Journal,* 22(2), pp. 82-96. https://doi.org/10.62273/OWZR5858

Analytics for an Audience of Healthcare Professionals: Curriculum Design and Student Perceptions

Jennifer Xu and Monica Garfield

1. INTRODUCTION

Analytics is one of the two primary thematic areas in which the information systems (IS) discipline can offer tremendous help to transform healthcare through research and education (Kohli & Tan, 2016). Healthcare analytics can be used to enhance patient care, increase quality of service, reduce costs and medical errors, and improve patient satisfaction (Strome, 2013). Recent years have seen a rapidly increasing demand for healthcare analytics by hospitals and medical institutions, disease control centers, insurance companies, and other healthcare organizations (Bates et al., 2014; Zhang, 2018). Several driving forces have contributed to this trend: the need to reduce cost of care (Bates et al., 2014; Osawa et al., 2020); the availability of big health data in the form of Electronic Health Records (EHRs) and multimedia data generated by mobile devices, monitors and sensors, and social media users (Dolezel & McLeod, 2019); as well as the advancements in machine learning (ML) and artificial intelligence (AI) technologies that have enabled deep analyses of large volumes of data to discover novel patterns and knowledge (Rajkomar et al., 2019; Rajpurkar et al., 2022; Yang et al., 2021).

In response to the trend, many universities have established programs and courses with a focus on healthcare analytics (Paul & MacDonald, 2020). However, most of these courses are designed for undergraduate or graduate students who are enrolled in IS, information technology (IT), or analytics majors, aiming at careers in data science and analytics in the healthcare industry. Little pedagogical research and guidance can be found regarding the development of analytics curricula for healthcare professionals (e.g., doctors, nurses, and hospital managers), who are enrolled in non-analytics, non-IS programs (e.g., executive MBA programs) seeking to understand and use analytics to help with decision making in clinical or administrative tasks.

Healthcare professionals in such programs may be different from IS/IT and analytics students in two aspects, posing challenges on the curriculum design. First, unlike IS/IT and analytics students who often have taken some prerequisite technical courses (e.g., data processing and management courses), healthcare professionals may not necessarily have technical background and skills, especially when healthcare analytics is the only technical/analytical course in the program. Moreover, while some professionals may have received training in statistics and regression, which is fundamental to analytics, others may not be adequately prepared for data analysis. Consequently, it is difficult to determine the appropriate scope, *topic coverage*, and pacing in such a situation, if allocating introductory and advanced contents into two courses is not an option.

Second, healthcare professionals may use analytics for different purposes due to their diverse job roles and responsibilities in their organization. For instance, while doctors may hope to be able to explain to patients how algorithms work to make diagnoses, managers may have little interest in clinical analytics but only wish to visualize the organization's operational expense data. With the varying goals, it is hard to consolidate a set of *learning objectives* that meet everyone's expectations.

To address these challenges, this paper presents the design of an executive MBA course on healthcare analytics at the business school of a northeastern U.S. university. The course is one of the required courses of the program specifically oriented for healthcare professionals. We also report the feedback from the students who are employees of a world-class hospital based in the Greater Boston area. We seek to address two research questions (RQs):

- **RQ1**: What *topics* should be covered, and *software tools* be used in the healthcare analytics course?
- **RQ2**: How do students *perceive* the effectiveness of various aspects of the course (e.g., topics, tools, learning activities, assessment methods, and the fulfillment of learning objectives)?

The contribution of this research is threefold: First, this study presents the curriculum design of a healthcare analytics course for professionals enrolled in non-analytics, non-IS, executive programs. Educators seeking to design analytics courses for similar audiences may find the content topics, selected tools, and assessment methods helpful for their curriculum development endeavors. Second, we summarize lessons learned from this course and offer a few design principles. Third, the students' feedback reveals additional aspects that course designers and program developers need to consider carefully when facing a professional audience with diverse learning objectives and expectations.

The remainder of this paper is organized as follows. The next section reviews the related work. Section 3 presents the research methodology, followed by the course design components. Section 5 reports the results of the survey. Section 6 discusses the results, lessons learned, and future directions. The last section concludes the paper.

2. RELATED WORK

Healthcare Analytics

Healthcare analytics is a broad term referring to a collection of methods, tools, and techniques to "explore, analyze, and extract value and insight from healthcare data" (Strome, 2013, p. 2). One of the reasons for the increasing adoption of healthcare analytics is the fast-growing cost of healthcare. According to the World Health Organization (2023), the United States has been the most expensive country in healthcare with total healthcare expenditure counting for approximately 17% of its Gross Domestic Product (GDP). Another driving force is the wide adoption of Electronic Health Record (EHR) systems, which makes it possible to analyze medical records and histories of patients in a timely manner to identify risk factors and design effective treatment procedures (Kohli & Tan, 2016). Big data, such as mobile device and sensor data, medical imaging data, and text data, have also been analyzed and mined in order to assist with disease control and public health by monitoring the spread of viruses and infectious diseases (Xu & Babaian, 2022).

State-of-the-art analytics techniques can be used to reduce healthcare cost by identifying high-risk, high-cost patients (Bates et al., 2014; Osawa et al., 2020), reducing preventable deaths and medical errors (Corny et al., 2020), and optimizing clinical workflow (Akkus et al., 2019). In addition to traditional statistical modeling and data mining techniques such as regression analysis and decision trees, techniques based on machine learning (ML) and artificial intelligence (AI) has been widely employed to aid descriptive, predictive, and prescriptive analysis in various clinical areas (Rajkomar et al., 2019), such as prediction of diabetes and hypertension (Razavian et al., 2015), diagnosis of cardiovascular diseases (Litjens et al., 2019), and detection of cancerous tumors (Lehman et al., 2019).

Healthcare organizations, providers, and patients could benefit significantly from understanding and effectively using healthcare analytics. Many healthcare professionals wish to receive education and training from properly designed programs and courses to gain knowledge and skills, and to learn to understand, select, and use state-of-the-art analytics technologies to enhance the quality of care and reduce healthcare costs.

Curriculum Design of Healthcare Analytics

Although curriculum frameworks for healthcare informatics have been recommended by practitioners (e.g., HIMSS (2024) and CAHIIM (2024)), only a limited number of studies can be found in the literature that present the curriculum design of healthcare analytics, and the subjects and contents in these courses are rather diverse. Such curricula research typically focuses on three key issues: content topics, tools, and skill sets. For example, Dolezel and McLeod (2019) conducted a comprehensive survey with healthcare managers across the nation. They found that data mining, data visualization, and SQL are the most frequently used technologies in healthcare workplaces, and that analytics tools including SAS, IBM SPSS, Tableau, and Microsoft Power BI are the top visualization tools adopted. Paul and MacDonald (2020) presented a curriculum framework for creating new or evaluating existing analytics programs at both undergraduate and graduate levels. They identified a set of required skills (e.g., communication skills and business acumen) and recommended to incorporate into a healthcare analytics program a sequence of courses ranging from basic statistics to advanced machine learning. Zheng et al. (2014) considered the broader healthcare IT (HIT) programs and suggested bringing business intelligence and analytics components to the HIT curriculum in addition to other HIT courses such as data management, data warehousing, and EHR systems. Parks (2020) proposed to employ a contextual active learning approach to designing a healthcare analytics course in a business curriculum. By contextualizing pedagogical activities (e.g., lectures, class discussions, and assignments) in a sequence of modules related to the healthcare domain, this design approach helps better engage students and achieve the key learning objectives. Sapci and Sapci (2020) performed a systematic review of the HIT education literature and recommended integrating AI training into medical and health informatics curricula. They proposed to teach AI as a new competency and suggest three types of skills: application of AI techniques, development of AI applications, and assessments of AI limitations and validation of clinical accuracy of AI algorithms.

Nearly all these curricula are designed for IS/IT or analytics students who have already had some technical skills. It is unclear how these curricula would fit the needs of healthcare professionals who have little IT background, limited technical knowledge and skills, and diverse learning objectives. Therefore, the design of a course for this type of audience must take these differences into consideration when selecting the content topics, pedagogical and learning activities, software tools, and assessment methods.

3. METHODOLOGY

This course was part of an executive MBA program offered at a business school in a northeastern U.S. university. The program was designed specifically for healthcare professionals, and the first cohort of students came from a world-class hospital based in the Greater Boston area. Since our students had diverse job roles, expectations, and goals, the first step in our course development was to send inquiries to several student representatives to gather learning objectives (LOs), which will be presented in the next section.

To address the first research question (RQ1) regarding content topics, we did a textbook search. Unfortunately, as noted in previous research (Parks, 2020), there had been no suitable textbooks on healthcare analytics. There also was no pedagogical framework or curriculum model to follow. As a result, we compiled a set of lecture notes, based on the identified LOs (see the next section), using materials from diverse sources including data mining and business analytics textbooks (e.g., (Shmueli et al., 2016)), academic literature databases (e.g., PubMed Central, Google Scholar), and online resources (e.g., GitHub).
To answer the second research question (RQ2) regarding student perceptions of the course design, we used the survey methodology to gather their feedback. The survey was anonymous and consisted of two parts: Part A was administered in the first week of the semester and was intended to gather information about student backgrounds, demographics (e.g., age, gender), job roles, prior analytics knowledge, and learning expectations. Part B was administered in the last week of the semester and focused on student perceptions of the effectiveness of the different aspects of the course design, measured by 5-point Likert scale questions with 1 being the least favorable and 5 the most favorable option.

Questions in Part B were developed based on previous research. In particular, the ease-of-use, usefulness, and satisfaction of the course design component (e.g., topics, tools) were assessed using instruments found in technology acceptance model (TAM) research (Davis 1985). A pilot survey was conducted before the last week to validate the questions.

Appendix A provides the complete questionnaires of the two parts.

4. COURSE DESIGN

This section presents the course design regarding the learning objectives, topic coverage, tools selected, and assessments of learning outcomes.

Learning Objectives
Based on the responses from representative students (see the above section), we identified four key LOs for this course:

- LO1: To be able to *select, process, and visualize* healthcare data that are appropriate for an analytics task.
- LO2: To *understand* the methods and algorithms conceptually, and to be able to *select and use* the appropriate methods and algorithms based on the requirements of the task.
- LO3: To be able to *interpret* the results produced by algorithms and tools in the healthcare context and make decisions based on the results.
- LO4: To be cautious of the ethical issues of using healthcare analytics in decision support.

Various pedagogical and learning activities were used in this course, including lectures, in-class hands-on exercises, case studies, and class discussions.

The course was delivered as a synchronous online course on Zoom.

Week	Module	Topics	Case Studies
1	1. Basic Concepts	• Components in healthcare analytics applications • Use cases of healthcare analytics	(Bates et al., 2014; Levin et al., 2018; Morel et al., 2020; Osawa et al., 2020)
2-4	2. Data Processing and Visualization	• Healthcare data sources and data types • Data quality and data imputation • Data visualization	(Bhaskaran & Smeeth, 2014; Kahn et al., 2012; Miao et al., 2023; Weiskopf & Weng, 2013)
5-7	3. Statistical Analysis	• ANOVA • Linear regression • Logistic regression	(Bhandari et al., 2020; Skrepnek, 2005)
9-12	4. Machine Learning (ML) and Artificial Intelligence (AI)	• Decision tree • Bayesian models • Neural networks and deep learning • Natural language processing (NLP) • Image processing in radiology • Performance evaluation	(Akkus et al., 2019; Corny et al., 2020; Lehman et al., 2019; Litjens et al., 2019; Rajkomar et al., 2019; Yang et al., 2021)
13	5. Ethics of Healthcare Analytics	• Ethical and legal issues • Data privacy • Fairness and biases • Limitations of algorithms	(Balthazar et al., 2018; Burton et al., 2017; Cohen et al., 2014; Katznelson & Gerke, 2021)

Table 1: Topic coverage

Topic Coverage
Based on our review of the literature and business analytics textbooks, a set of topics were selected and organized into several topic modules: basic concepts of healthcare analytics, data processing and visualization, statistical analysis, machine learning (ML) and artificial intelligence (AI), and ethical issues. Each module consisted of one or more class meetings. Table 1 presents the topic coverage of this course.

Module 1: Basic concepts (LO1). This module provided an overview of healthcare analytics and introduced the basic concepts of analytics, such as descriptive, predictive, and prescriptive analysis. Five use cases in which healthcare analytics for quality of care and cost reduction were presented, including high-cost patients, readmission, triage, decompensation, and adverse events (Bates et al., 2014). Each use case was illustrated with a case study selected from the literature. For example, a case study on the identification of high-need, high-cost patients was used to demonstrate the potential of machine learning in reducing medical costs (Osawa et al., 2020).

Module 2: Data processing and visualization (LO1/LO3). This module introduced the diverse sources of healthcare data, including EHRs, sensor and mobile device data, images, text, videos, and audio data. Different types of data (e.g., continuous vs. categorical, structured vs. unstructured) were explained. Special attention was drawn to the data quality issue because poor data quality may lead to incorrect or misleading results causing devastating consequences to patients. Various data imputation approaches were presented for handling different types of missing data problems. A comprehensive case study was presented and discussed regarding how different data quality issues were resolved in a clinical study seeking to identify risk factors for 30-day readmission for hip fracture surgeries (Miao et al., 2023).

Students learned to use Tableau (see the next subsection on tools) to visualize healthcare data by creating various charts and dashboards. They were required to interpret the visualizations and to examine the validity of the results in the particular setting of the task based on their domain knowledge.

Module 3: Statistical analysis (LO2/LO3). Since statistics is the foundation of many analytical approaches, including machine learning (ML), this module is critical for students to understand and use any advanced techniques. Although students learned basic statistics previously, most of them were not familiar with regression. As a result, three class meetings were spent on reviewing the concepts and practicing statistical analysis using SPSS. Students learned when to use linear or logistic regression depending on whether the outcome variable was continuous or categorical, and how to interpret the results produced by the software and draw valid conclusions.

Module 4: ML and AI (LO2/LO3). This was the advanced topic module of this course. This module focused on supervised learning (i.e., classification) and introduced not only traditional ML approaches such as decision trees and Bayesian methods, but also the most up-to-date advancements in ML and AI, such as deep learning and its applications in medical image processing and natural language processing (NLP). Given the background characteristics of the students, the lectures did not put a heavy weight on the inner workings of the algorithms (e.g., back propagation in deep learning) but only gave brief conceptual descriptions of the logic underlying different algorithms. A number of case studies, such as diabetes prediction using neural networks (Razavian et al., 2015) and symptom recognition from patient narratives (Xu & Babaian, 2022), were discussed in class to demonstrate the applications of ML and AI in healthcare.

Students also learned to select appropriate algorithms based on the task and algorithm performance evaluation metrics (e.g., accuracy, sensitivity, specificity, false positive/negative rate, precision, recall, and F1 score).

Module 5: Ethics of healthcare analytics (LO4). The use of analytics for decision-making in the healthcare settings may raise various ethical and legal challenges and risks, especially when a decision is a critical one, such as diagnoses of diseases, selection of treatment options, and allocation of medical resources (Cohen et al., 2014). This module took a seminar format, in which students discussed and debated on many ethical issues based on their own observations and experiences from work.

Software Tools

Since most students did not have programming skills, we selected software tools that required no programming, including Microsoft Excel for basic data processing, Tableau for data visualization, IBM SPSS for statistical analysis, and RapidMiner for ML. These tools were quite easy to learn and use. With a few demonstrations and hands-on exercises, students became familiar with the tools and felt comfortable using them.

Assessments of Learning Outcomes

Learning outcomes were assessed using homework assignments, a midterm exam, a final project, and class participation, which accounted for 30%, 30%, 30%, and 10% of the grade, respectively.

Four assignments were given for students to practice data visualization (Module 2), statistical analysis (Module 3), traditional ML (Module 4), and deep learning (Module 4), respectively. For each assignment, students were provided with one or more healthcare datasets retrieved from public sources (e.g., Kaggle.com). Given the students' busy work schedules, each assignment was kept at an appropriate length and difficulty level so that it could be completed within 30-45 minutes.

The midterm exam was administered in week 8, and consisted of multiple-choice, short-answer, and essay questions. For the final project, students were given the option to either complete it independently or collaborate with another classmate. Students were required to select a healthcare problem they would like to study (e.g., youth obesity); acquire the datasets either from public or proprietary sources; process, visualize, and analyze the data using at least two methods; write the report and present the project in the last week.

In the last week of the semester, Part B of the survey was administered to gather student feedback.

5. SURVEY RESULTS

Student Backgrounds

Twenty-seven students who enrolled in this course participated in both parts of the survey. The sample included 5 (18.5%) males, 20 (74.1%) females, and 2 (7.4%) people who chose not to disclose gender information. In terms of age, 4 (14.8%) students were in their 20s, 7 (25.9%) in 30s, 11 (40.7%) in 40s, 3 (11.1%) in 50s, and 2 (7.4%) people chose not to disclose age information. Approximately half of the students (n=14, 51.9%) had bachelor's degrees, 11 (40.7%) had master's degrees, and 2 (7.4%) students had doctoral degrees. Most students (n=18, 66.7%) had administrative job roles in their organization, and a small number of students had other roles including clinical (n=4, 14.8%), technical (n=1, 3.7%), and support (n=2, 7.4%). Their length of tenure with the current role ranged from less than one year (n=5, 18.5%), 1-9 years (n=18, 66.7%), 10-15 years (n=1, 3.7%), to 16 or more years (n=3, 11.1%).

Nearly all students (n=26, 96.3%) took basic statistics courses previously. However, only three students knew regression analysis, and only a few students had IT knowledge or training in such subjects as database management (n=3, 11.1%), programming language (n=3, 11.1%), or analytics (n=1, 3.7%). Only one student used other analytical methods (e.g., decision trees) in addition to regression before.

Students had different learning objectives and expectations, including learning to interpret results generated by tools (n=21, 77.5%), to understand how algorithms work (n=20, 74.1%), to choose the appropriate analytics methods (n=14, 51.9%), to use the state-of-the-art methods and tools (n=14, 51.9%), to understand the limitations of different methods (n=12, 44.4%), and to increase career opportunities (n=11, 40.7%).

Perceptions of Course Design

Part B of the survey concerned students' perceptions of the course design in terms of topic coverage, learning activities, software tools, assessment methods, and fulfillment of learning objectives. We also measured their satisfaction with the course and solicited their qualitative comments.

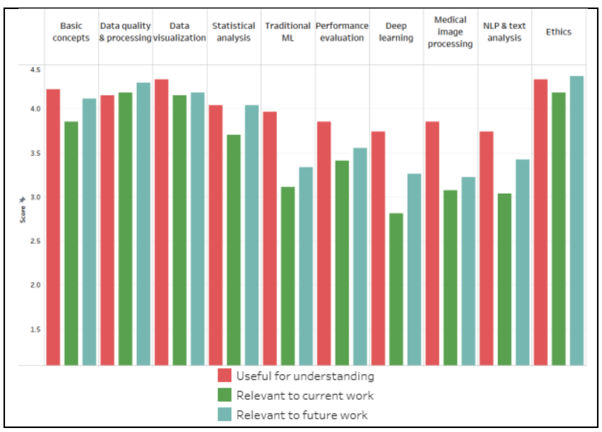

Figure 1. Topic usefulness and relevance

Topic coverage. Figure 1 presents students' perceptions of the 10 major topics in term of their *usefulness* for understanding healthcare analytics in general and *relevance* to their current and future work. It shows that students considered data visualization and the discussions about ethics the most useful topics. The average score for both topics was 4.33 out of 5.0, followed by 4.22 for basic concepts, 4.15 for data quality and processing, and 4.04 for statistical analysis. ML topics including deep learning, image processing, and NLP were considered the least useful topics with average scores below 4.0. Similar patterns are found in perceptions of relevance with ethics, data quality, and visualization being rated the most relevant and ML topics the least relevant.

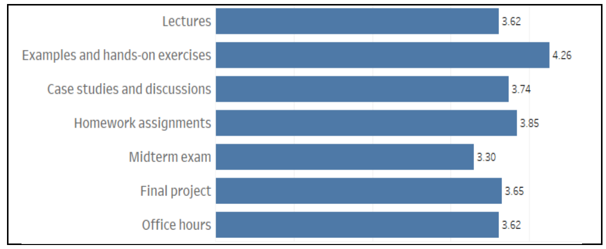

Figure 2. Usefulness of learning activities

Learning activities. Figure 2 displays student perceptions of the usefulness of various learning activities including lectures, hands-on exercises, case studies and discussions, homework assignments, midterm exam, final project, and office hours. It appears that students liked examples and hands-on exercises in class (4.26) and considered them useful for understanding the concepts and materials. The next three useful activities are homework assignments (3.85), case studies and class discussions (3.74), and lectures (3.62). They perceived the midterm exam as least useful (3.30).

Figure 3. Perceptions of the tools

Software Tools. Figure 3 presents the scores for the perceptions of software tools regarding their usefulness and ease-to-use, as well as the likelihood of using them in students' future work. Excel clearly was perceived as the most useful and easy-to-use tool likely because students all had been very familiar with Excel prior to the class. The runner up was Tableau, followed by SPSS. RapidMiner was rated the least useful and easy-to-use and also scored the lowest for future use. This could have been caused by the fact that it was only used for ML, which was perceived as least relevant by students.

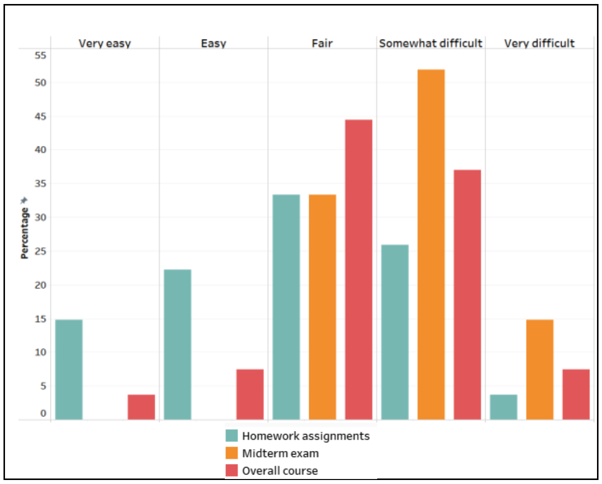

Figure 4. Perceptions of the difficulty levels

Assessments. Figure 4 shows the percentages of students with their perceptions about the difficulty levels of course work (i.e., homework assignments and the midterm exam), ranging from being very easy to very difficult. Since students had not completed their final projects by the time of the survey, we did not ask about their perceptions of the final project, but instead asked their overall impressions of the difficulty level of this course.

It appears that the distribution of the homework assignments is tilted toward the easy side, and that for the midterm is toward the hard side. More than half of the students thought the homework assignments were relatively easy or fair, 26% of the students had some difficulties with the assignments, and 3.7% found them very difficult. In contrast, only one third of the students thought the midterm was fair, and the rest of them thought it was somewhat difficult (51.2%) or very difficult (14.8%). The distribution of the overall perceptions of the course is centered on the fair (44.4%) level and the somewhat difficult level (37.0%).

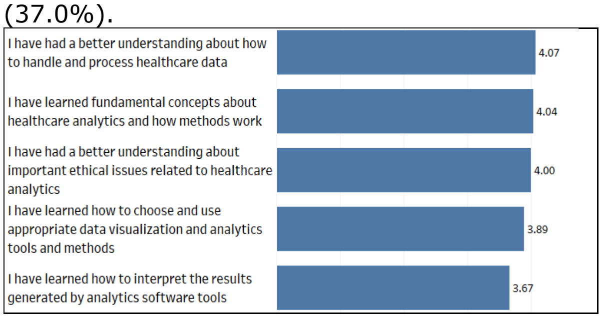

Figure 5. Fulfillment of learning objectives

Learning objectives. Figure 5 presents students' scores regarding how much they perceived that the course fulfilled their learning objectives.

Satisfactions. The average scores for the overall satisfaction of the course and the likelihood of recommending the course to others are both 4.0 out of 5.0.

Qualitative comments. Students made many comments regarding the strengths of the course and also suggested ways to improve the course design. For example, a student commented on the roles of different learning activities:

> *The homework assignments were excellent opportunities to implement what we were learning in class. The real-life examples and case studies of how someone uses these concepts professionally were very helpful. The final project is a great way to tie everything together.*

Most students enjoyed the data visualization part of the course in which they learned to use Tableau. A student put it enthusiastically:

> *I loved the Tableau dashboards and am looking to create these for a few KPIs in our billing and operations.*

A student made a very useful recommendation about how to improve the future delivery of the course by balancing the time spent on clinical and other data analyses:

> *I would suggest less of a focus on exactly how data is used in clinical care by clinicians and more on the healthcare data world as a whole. Many of us are not patient facing and that element (the actual patient care/impact) feels like a small part of the chain of data analytics.*

6. DISCUSSION

Several findings can be drawn from the survey results. First, in terms of *topic coverage*, students enjoyed data visualization the most and considered advanced topics on ML and AI less useful and relevant. This is not surprising because ML and AI have been mainly used in advanced clinical applications such as disease predictions and most students may not need these advanced techniques to analyze their data. Students also liked the discussion about ethics. Second, similar to the perceptions of topics, the perceptions of *software tools* favored Excel and Tableau, yet SPSS and RapidMiner, which were used for advanced statistical analysis (e.g., logistic regression) and ML, were perceived as less useful and difficult to use. Third, regarding *learning activities and assessments*, students preferred hands-on practices, case studies, and homework assignments, but did not like the timed midterm exam. Last, students found the course moderately difficult overall, and fulfilled most of their *learning objectives* and expectations.

Based on these findings, we summarize a few lessons learned:

- *The topic coverage should be determined based on the students' background and learning expectations*. Since this course is the only analytics course in the program, it could be too ambitious to treat introductory topics (e.g., data processing and visualization) and advanced techniques (e.g., ML and AI) equally. Students had more difficulty understanding and appreciating the logic and power of ML and AI methods, causing them to consider these topics less useful. Future deliveries of this course could allocate more class meetings to introductory topics and use just fewer classes to explore ML and AI techniques at the conceptual level.

- *The course could leverage both simple and advanced analytics tools*. Simple data analysis functionality, such as descriptive statistics, ANOVA, and linear regression, and visualization tools, such as charts and diagrams, are readily available in Microsoft Excel. Since students were already familiar with Excel, they found Excel useful and easy to use. They also enjoyed Tableau which allowed them to visualize their data using various charts and dashboards. However, they tended to find SPSS and RapidMiner intimidating due to the relatively steep learning curves for these tools. Corresponding to the topic coverage adjustments (see above), we would increase the time and use of Tableau and even other popular data visualization tools, such as Microsoft Power BI. However, we will still spend a reasonable amount of time on advanced tools such as SPSS and RapidMiner, which may be challenging for some students, to bring up the full potential of healthcare analytics.

- *Considering students' characteristics, the assessment methods need to be adjusted.* Since students were professionals in an executive education program, they may not be as accustomed to test taking as regular students. As a result, they tended to dislike typical assessment methods such as timed exams. Alternative methods, such as a take-home midterm exam may serve this audience better.

- *The course should cover a balanced set of applications of both clinical and administrative analytics*, which use different types of data and serve different purposes. Clinical applications by doctors and researchers need to use patient health records in order to identify risk factors and design appropriate intervention measures. However, hospital managers analyze operational, financial, and admission data for quality control (e.g., to identify factors that affect 30-day readmission rates) or pricing and revenue analysis purposes. With different job roles and expectations, students would need to learn different types of data analysis and applications. As noted by one student, if most of the students in the cohort had administrative roles, less time could be allocated to clinical analytics.

7. CONCLUSION

Based upon our thorough review of currently published work on the subject of curriculum design for healthcare analytics, we find that our research is unique in its focus on healthcare analytics education for professional audiences. The design aspects that we share in this paper, including topics, tools, and assessments, would be helpful for other educators to design similar courses or programs intended for healthcare professionals. In addition, the perceptions of the students and the lessons learned reveal additional aspects, expectations, attitudes toward such a course that educators may need to consider when designing courses for professional audiences.

In addition to the modifications that we will implement in the future delivery of this course, we will perform additional empirical studies to investigate how different course design factors affect student performance, perceptions, and satisfaction in this course.

8. REFERENCES

Akkus, Z., Cai, J., Boonrod, A., et al. (2019). A survey of deep-learning applications in ultrasound: Artificial intelligence–powered ultrasound for improving clinical workflow. *Journal of the American College of Radiology, 16*(9),1318-1328. https://doi.org/10.1016/j.jacr.2019.06.004.

Balthazar, P., Harri, P., Prater, A., et al. (2018). Protecting your patients' interests in the era of big data, artificial intelligence, and predictive analytics. *Journal of the American College of*

Radiology, 15(3, Part B), 580-586. https://doi.org/10.1016/j.jacr.2017.11.035.

Bates, D. W., Saria, S., Ohno-Machado, L., et al. (2014). Big data in health care: Using analytics to identify and manage high-risk and high-cost patients. *Health Affairs, 33*(7), 1123-1131. https://doi.org/10.1377/hlthaff.2014.0041.

Bhandari, S., Shaktawat, A., Tak, A., et al. (2020). Logistic regression analysis to predict mortality risk in COVID-19 patients from routine hematologic parameters. *Ibnosina Journal of Medicine and Biomedical Sciences, 12*(2), 123-129. https://doi.org/10.4103/ijmbs.ijmbs_58_20.

Bhaskaran, K., & Smeeth, L. (2014). What is the difference between missing completely at random and missing at random? *International Journal of Epidemiology, 43*(4), 1336-1339.https://doi.org/10.1093/ije/dyu080.

Burton, E., Goldsmith, J., Koenig, s., et al. (2017). Ethical considerations in artificial intelligence courses. *AI Magazine, 38*(2), 22-34. https://doi.org/10.1609/aimag.v38i2.2731.

CAHIIM (2024). HI Curriculum Requirements. https://www.cahiim.org/accreditation/health-informatics/curriculum-requirements.

Clayton, P. R., & Clopton, J. (2019). Business curriculum redesign: Integrating data analytics. *Journal of Education for Business, 94*, 57-63. https://doi.org/10.1080/08832323.2018.1502142.

Cohen, I. G., Amarasingham, R., Shah, A., et al. (2014). The legal and ethical concerns that arise from using complex predictive analytics in health care. *Health Affairs, 33*(7), 1139-1147. https://doi.org/ 10.1377/hlthaff.2014.0048.

Corny, J., Rajkumar, A., Martin, O., et al. (2020). A machine learning–based clinical decision support system to identify prescriptions with a high risk of medication error. *Journal of the American Medical Informatics Association, 27*(11), 1688–1694. https://doi.org/10.1093/jamia/ocaa154.

Davis, Fred D. (1985). A technology acceptance model for empirically testing new end-user information systems: Theory and results. Doctoral dissertation of Massachusetts Institute of Technology, 1985.

Dolezel, D., & McLeod, A. (2019). Big data analytics in healthcare: Investigating the diffusion of innovation. *Perspectives in Health Information Management, 16*(Summer), 1a.

HIMSS (2024). Health Informatics. https://www.himss.org/resources/health-informatics.

Kahn, M. G., Raebel, M. A., Glanz, J. M., et al. (2012). A pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research. *Medical care, 50*(0). https://doi.org/10.1097/MLR.1090b1013e318257dd318267.

Katznelson, G., & Gerke, S. (2021). The need for health AI ethics in medical school education. *Advances in Health Sciences Education, 26*(4), 1447-1458. https://doi.org/10.1007/s10459-021-10040-3.

Kohli, R., & Tan, S. S.-L. (2016). Electronic health records: How can IS researchers contribute to transforming healthcare? *MIS Quarterly, 40*(3), 553-573. https://doi.org/10.25300/MISQ/2016/40.3.02/

Lehman, C. D., Yala, A., Schuster, T., et al. (2019). Mammographic breast density assessment using deep learning: Clinical implementation. *Radiology, 290*(1), 52-58. https://doi.org/10.1148/radiol.2018180694.

Levin, S., Toerper, M., Hamrock, E., et al. (2018). Machine-learning-based electronic triage more accurately differentiates patients with respect to clinical outcomes compared with the emergency severity index. *Annals of Emergency Medicine, 71*(5), 565-574.e562. https://doi.org/10.1016/j.annemergmed.2017.08.005.

Litjens, G., Ciompi, F., Wolterink, J. M., et al. (2019). State-of-the-art deep learning in cardiovascular image analysis. *JACC: Cardiovascular Imaging, 12*(8_Part_1), 1549-1565.

Miao, Z., Sealey, M. D., Sathyanarayanan, S., et al. (2023). A data preparation framework for cleaning electronic health records and assessing cleaning outcomes for secondary analysis. *Information Systems, 111*, 102130. https://doi.org/10.1016/j.is.2022.102130.

Morel, D., Yu, K. C., Liu-Ferrara, A., et al. (2020). Predicting hospital readmission in patients with mental or substance use disorders: A machine learning approach. *International Journal of Medical Informatics, 139*, 104136. https://doi.org/10.1016/j.ijmedinf.2020.104136.

Osawa, I., Goto, T., Yamamoto, Y., et al. (2020). Machine-learning-based prediction models for high-need high-cost patients using nationwide clinical and claims data. *npj Digital Medicine,* *3*(1), 1-9. https://doi.org/10.1038/s41746-020-00354-8.

Parks, R. F. (2020). A pedagogic experience in designing a healthcare analytics course: Lessons learned. *Information Systems Education Journal, 18*(5), 4-15.

Paul, J. A., & MacDonald, L. (2020). Analytics curriculum for undergraduate and graduate students. *Decision Sciences Journal of Innovative Education, 18*(1), 22-58. https://doi.org/10.1111/dsji.12196.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine, 380*(14), 1347-1358. https://doi.org/10.1056/NEJMra1814259.

Rajpurkar, P., Chen, E., Banerjee, O., et al. (2022). AI in health and medicine. *Nature Medicine, 28*(1), 31-38. https://doi.org/10.1038/s41591-021-01614-0.

Razavian, N., Blecker, S., Schmidt, A. M., et al. (2015). Population-Level Prediction of Type 2 Diabetes From Claims Data and Analysis of Risk Factors | Big Data. *Big Data, 3*(4), 277-287. https://doi.org/10.1089/big.2015.0020.

Sapci, A. H., & Sapci, H. A. (2020). Teaching hands-on informatics skills to future health informaticians: A competency framework proposal and analysis of health care informatics curricula. *JMIR Medical Informatics, 8*(1), e15748. https://doi.org/10.2196/15748.

Shmueli, G., Bruce, P. C., & Patel, N. R. (2016). *Data Mining for Business Analytics*. Hoboken, NJ: Wiley.

Skrepnek, G. H. (2005). Regression Methods in the Empiric Analysis of Health Care Data: Journal of Managed Care Pharmacy. *Journal of Managed Care + Specalty Pharmacy, 11*(3), 240-251. https://doi.org/10.18553/jmcp.2005.11.3.240

Strome, T. L. (2013). *Healthcare Analytics for Quality and Performance Improvement*. Hoboken, NJ: Wiley.

Weiskopf, N. G., & Weng, C. (2013). Methods and dimensions of electronic health record data quality assessment: Enabling reuse for clinical research. *Journal of the American Medical Informatics Association, 20*(1), 144-151. https://doi.org/10.1136/amiajnl-2011-000681.

World Health Organization (2023). *Global Health Expenditure Database*. https://apps.who.int/nha/database/Select/Indicators/en.

Xu, J., & Babaian, T. (2022). Investigating generalizability of BERT models for symptom recognition. In *Proceedings of the 32nd Workshop on Information Technologies and Systems (WITS)*, Copenhagen, Denmark.Yang, S., Zhu, F., Ling, X., et al. (2021). Intelligent health care: Applications of deep learning in computational medicine. *Frontiers in Genetics, 12*. https://doi.org/10.3389/fgene.2021.607471.

Zhang, C. (2018). Using open healthcare datasets to teach health data analytics. *AMCIS 2018 Proceedings*.

Zheng, G., Zheng, C., & Li, L. (2014). Bringing business intelligence to health information technology curriculum. *Journal of Information Systems Education, 25*(4), 317-326.

APPENDIX A

Survey Items

Attributes	Items
Age	What is your age? ○ 20-29 ○ 40-49 ○ 60+ ○ 30-39 ○ 50-59 ○ Prefer not to answer
Gender	What is your gender? ○ Male ○ Female ○ Prefer not to answer
Education	What is the highest degree you have completed? ○ Associate degree ○ Bachelor's degree ○ Master's degree ○ Doctorate degree
Job Role	What is your job role (check all that apply) in your organization? ☐ Clinical (e.g., Physician, Nurse, Surgeon) ☐ Administrative (e.g., Director, Manager, Team Leader) ☐ Technical (e.g., Radiology) ☐ Support (e.g., Educator, Analyst) ☐ Other, please specify_______________
Tenure	How long have been in your current job role? ○ Less than one year ○ 1-5 years ○ 6-9 years ○ 10-15 years ○ 15+ years
Technical Background	Have you taken any of the following courses at any institution (check all that apply)? ☐ Statistics ☐ Data science or analytics ☐ Data management ☐ Programming languages (e.g., Python, R, Java) ☐ Machine learning and artificial intelligence ☐ Other, please specify_______________
Analytical Knowledge	Have you used any of the following analytical methods in your work (check all that apply)? ☐ Regression analysis (e.g., Linear, Logistic) ☐ Bayesian models (e.g., Naïve Bayes, Belief Network) ☐ Decision tree ☐ Support vector machine (SVM) ☐ Neural networks ☐ Natural language processing (NLP) ☐ Clustering ☐ Other, please specify_______________

Learning Expectations	What objectives do you wish to achieve in this course? ☐ I'd like to learn the state-of-the-art methods and tools ☐ I'd like to understand how analytical methods work ☐ I'd like to be able to understand and interpret the resulted generated by analytical tools ☐ I'd like to learn limitations of different methods ☐ I'd like to be able to choose the appropriate methods when analyzing my data ☐ I want to increase my career opportunities ☐ Other, please specify_______________

Table 1. Survey items in Part A: Student Background

Course Aspects	Items	Measures (5-point Likert Scale)
Topics: Usefulness	Do you think that the topics are useful for you to understand healthcare analytics in general? - Basic concepts - Data quality and processing - Data visualization - Statistical analysis - Traditional ML (e.g., Decision tree, Bayesian) - Performance evaluation - Deep learning - Medical image processing - NLP & text analysis - Ethics	From useless (1) to extremely useful (5) for each topic
Topics: Relevance to Current Work	To what extent do you think the topics are relevant to your **current** work? - Basic concepts - Data quality and processing - Data visualization - Statistical analysis - Traditional ML (e.g., Decision tree, Bayesian) - Performance evaluation - Deep learning - Medical image processing - NLP & text analysis - Ethics	From completely irrelevant (1) to extremely relevant (5) for each topic
Topics: Relevance to Future Work	To what extent do you think the topics are relevant to your **future** work? - Basic concepts - Data quality and processing - Data visualization - Statistical analysis - Traditional ML (e.g., Decision tree, Bayesian) - Performance evaluation - Deep learning - Medical image processing - NLP & text analysis - Ethics	From completely irrelevant (1) to extremely relevant (5) for each topic
Tools: Usefulness	How do you rate the usefulness of the software tools we used in this course? - Excel - SPSS - Tableau - RapidMiner	From useless (1) to extremely useful (5) for each tool
Tools: Ease-of-use	How do you rate the ease-of-use of the software tools we used in this course? - Excel - SPSS - Tableau - RapidMiner	From very difficult (1) to very easy (5) for each tool

Tools: Likely to Use in the Future	How likely do you think you will use these tools in your future work? - Excel - SPSS - Tableau - RapidMiner	From not likely (1) to very likely (5) for each tool
Learning Activities: Usefulness	How do you think the following learning activities help you understand the materials? - Lectures - Examples and hands-on exercises - Case studies and discussions - Homework assignments - Midterm exam - Final project - Office hours	From useless (1) to extremely useful (5) for each activity
Difficulty Level: Homework	How do you rate the difficulty level of the homework assignments? - HW1: Data visualization - HW2: Statistical analysis - HW3: Traditional ML - HW4: Deep learning	From extremely easy (1) to extremely difficult (5) for each assignment
Difficulty Level: Midterm Exam	How do you rate the difficulty level of the midterm exam?	From extremely easy (1) to extremely difficult (5)
Difficulty Level: Course	How do you rate the overall difficulty level of the course?	From extremely easy (1) to extremely difficult (5)
Learning Outcomes	After taking this course, I feel that - I have had a better understanding about how to handle and process healthcare data - I have learned fundamental concepts about healthcare analytics and how methods work - I have had a better understanding about important ethical issues related to healthcare analytics - I have learned how to choose and use appropriate data visualization and analytics tools and methods - I have learned how to interpret the results generated by analytics software tools	From strongly disagree (1) to strongly agree (5) for each statement
Satisfaction	How satisfied are you with the course design? How likely will you recommend this course to others?	From very dissatisfied (or very unlikely) to very satisfied (or very likely) for each question

Table 2. Survey items in Part B: Student Perceptions