*International Journal of Education and Development using Information and Communication Technology (IJEDICT),* 2022, Vol. 18, Issue 1, pp. 213-235

# Improving Data Visualization Skills: A Curriculum Design

**Daniel Asamoah**
**Wright State University, USA**

## ABSTRACT

It is unclear whether the data visualization skills possessed by data-centric practitioners reflect recent advances in computing systems and visual design principles, and their role in generating relevant and insightful output to support decision-making. In this study, the author presents a curriculum designed for teaching basic and advanced concepts in data visualization. The author used the survey method to examine whether respondents perceived the pedagogic approach to have enhanced their data visualization competence. The findings indicate preference for a hands-on pedagogic approach to learning data visualization as compared to an entirely theory-based approach. Furthermore, different components of the pedagogic approach were reviewed. The respondents rated most components of the curriculum highly as well as the pedagogical structure. As a cognitive tool, the respondents recognized that visualization helps to generate deeper insights into data, and it supports effective decision-making.

**Keywords**: *data visualization; curriculum design; pedagogy; visual analytics; decision-making*

## INTRODUCTION

A picture is worth a thousand words, yet not all pictures are worth one’s attention. Data visualization techniques have been traditionally used for effective information and insights generation. In recent years, visualization techniques, methodologies, and applications have experienced profound advancements such that many businesses and organizations perceive them as a primary means of analyzing and interpreting data. For instance, given the ubiquitous use of computers and semiconductor technology, computing systems now possess the ability to generate highly immersive and interactive visualization dashboards and charts. Earlier research revealed that about 57% of businesses implemented some form of data visualization as a way of generating business insights (Stodder 2013).

In the "Big Data" era, visualization has become a critical component of Information Systems, and it supports decision-making networking (Erbacher & Forcht 2010), and an increasing number of job postings now emphasize data visualization as a basic requirement (Dong & Triche 2020). For example, in accounting, data visualization skills can be leveraged to overlay revenue numbers on a map to portray insights related to location-based performance. Hence, training business students in the use of data visualization for decision-making has become a vital issue, and sharing ideas and experience on how to do so could be beneficial. Data visualization presents a fundamentally intuitive perspective for generating and communicating insights from data. Earlier research shows that for disciplines other than science and math, data visualization could be a tool for understanding and presenting insights that would otherwise require advanced statistical, mathematical, and computing skills (Owen, Domik, Ebert, Kohlammer, Rushmeier, Santos & Weiskopf 2013).

Data visualization is an integral component of visual analytics, widely considered a critical skill for data-centric jobs, and is one of the most sought after job requirements (Ryan, Silver, Laramee, Ebert & Rhyne 2019). Since 2016, there has been a steady increase in the number of job postings that list data visualization as a required skill (Ryan *et al.* 2019). As a cognitive tool, data visualization has been shown to enhance performance in job functions such as data analytics (Näykki & Järvelä 2008). Other domains such as creative arts attest to the use of visualization tools as an effective cognitive tool to improve performance (Odewumi 2021) and retention (Panjwani, Micallef, Fenech & Toyama 2009). It is critical for collaborative knowledge construction and allows for optimal knowledge contribution by individuals during group decision-making (Näykki & Järvelä 2008). Although the benefits of data visualization for decision-making is well-documented (Nguyen, Gardner & Sheridan 2020), application in both academia and industry is lacking. This could be due to an insufficient number of well-trained professionals who can leverage visualization outputs for decision-making (Srinivasa, Arun, James & Zhu 2021).

Academia as a stakeholder is tasked with training competent professionals who can leverage visualization skills to explore and generate insights from data. Most data visualization courses are taught using two main approaches depending on the leaners' technical background. The first approach provides students with practical experience related to data visualization and teaches practical concepts using computer programming systems. In the second approach, students are trained in theoretical concepts and best practices using specialized software such as Tableau (Tableau 2020) or Power BI (Microsoft 2020). The former approach is common with technical disciplines in which students may be conversant with computer programming and are comfortable generating visualization outputs by writing computer programming code. The latter approach is common with leaners that are non-technical and more comfortable with packaged visualization software applications. Either approach may help improve students' skills with data visualization. However, the knowledge gained tends to be focused on either practical or theoretical knowledge, but not both. This could reduce an individual's ability to generate insights from data using sound theoretical principles and visualization techniques.

Following calls and practices from prior studies (Owen *et al.* 2013; Janvrin, Raschke & Dilla 2014; Ryan *et al.* 2019), the author designed a data visualization course that focuses on both the underlying theory and principles for visualization as well as emphasizing practical competence using a packaged data visualization application. The course design considers a multidisciplinary environment where the leaners typically have a mixed technical and non-technical background with students from Management Information Systems, Finance, Accounting, Management, Marketing, and Economics. The mixed pedagogical approach implemented is expected to be more appropriate since it takes advantage of students' mixed background in teaching a wide spectrum of data visualization concepts. A mixed student population could be beneficial in fostering the discussion of concepts with a multi-disciplinary perspective which enriches the pedagogical experience (Domik 2011). The course requires no programming experience. However, students need to have basic knowledge and skills in working with a packaged software product, including Tableau and Microsoft Excel. Other relevant data visualization software and tools have been provided in Table B1 in Appendix C.

This study presents a pedagogical experience in teaching data visualization skills at a mid-western university in the Unites States. The pedagogical approach is based on a multi-year offering of a data visualization course to multi-disciplinary student groups. Using both statistical analysis and text mining, the author evaluates the effectiveness of the pedagogical approach on students' data visualization competency. The paper examines and reports relevant experiences in designing and teaching a data visualization course to both undergraduate and graduate students with diverse business backgrounds, including Accounting, Information Systems, and Finance. A parsimonious paradigm (Asamoah, Sharda, Hassan Zadeh & Kalgotra 2017) is used to present both technical and non-technical modules and assess the use of such modules in enhancing data visualization learning outcomes using a survey method. This study's main objective is to explore an effective curriculum that would support the teaching of data visualization to a multi-disciplinary student population. The primary research question asks if the curriculum design and pedagogic approach helps improve students' perceived data visualization competency.

The rest of the paper is designed as follows: an exploration of the previous literature about data visualization and how it has been taught to different leaners; then the paper delves into the curriculum design and the approach used in creating a comprehensive, yet parsimonious data visualization course targeted towards multi-disciplinary leaners with both technical and non-technical backgrounds. Next, the method for assessing the effectiveness of the pedagogic approach is discussed and the results are presented and discussed. The paper concludes with a summary and limitations of the study.

## LITERATURE REVIEW

Data visualization is defined as the communication of information using graphical presentations (Ward, Grinstein & Keim 2010). Using data visualization to analyze data and present valuable insights is not only vital to engineering and computer science disciplines, but also to business disciplines. For instance, disciplines such as finance and accounting utilize visualization techniques to analyze and present financial and accounting data in ways that are easily digestible to stakeholders. Prior research (Adams, Boersema & Mijksenaar 2010) shows that difficult concepts such as warning messages can be effectively portrayed when graphical design principles are used. Also, data visualization helps improve the analytics skills of business professionals with non-computing backgrounds and enhances their analytics capabilities (Asamoah *et al.* 2017; Nguyen, Gardner & Sheridan 2020). Furthermore, data visualization concepts have been found to be relevant in all disciplines for presenting business insights and knowledge, specifically in the area of business intelligence (Wixom, Ariyachandra, Goul, Gray, Kulkarni & Phillips-Wren 2011; Janvrin, Raschke & Dilla 2014).

Concepts in data visualization are taught in two main categories; Interactive Data Visualization (IDV) and Static Data Visualization (SDV) (Janvrin, Raschke & Dilla 2014). For IDVs, dashboards are used to visually present data while allowing end users to decide on what insights to portray and how to portray them. In SDVs, visuals are created by professionals who make decisions on what insights should be presented to end users and how they should be presented. Whereas IDV tends to be more effective in presenting a multi-faceted data, an efficient data analyst would rely on fundamental knowledge of SDV concepts. Hence, it is imperative that a data visualization course encompasses theory and practice from both IDV and SDV.

Prior research has identified three primary areas for data visualization (Owen *et al.* 2013). These are:

- Scientific or data visualization, mostly in the physical and biological sciences.
- Information visualization, mostly in business domains such as accounting and finance.
- Visual analytics, where data visualization is applied to data analytics. This uses data from multiple sources and with multiple formats.

The course that was developed leverages the third approach and focuses on the generation of insights and decision-making, using charts and other visualization techniques based on underlying business data. Previous studies have addressed multiple approaches and methodologies for teaching data analytics, business intelligence and data science courses (Gupta, Goul & Dinter 2015; Asamoah *et al.* 2017; Miah, Solomonides & Gammack 2020). These studies mainly focus on how to teach descriptive, predictive, and prescriptive analytics concepts in ways that enhance students' learning outcomes. For both research and practice, data visualization is mainly used to portray information during two phases of the analytics process. The first phase is during the data exploration phase where descriptive analytics is used to initially identify trends, depict anomalies, and show relationships between variables. In the second phase, data visualization is used to efficiently represent and interpret generated results. Most data visualization courses focus on the first phase, possibly in an effort not to overwhelm students with too much content. The result is that students become competent in the technical aspects of using visualization tools and applications to generate visual content, yet they find it challenging to interpret visualized content including related statistical concepts (Stenliden, Bodén & Nissen 2019).

Case studies are one of the most common methods for teaching data visualization concepts. For instance, in the case of Janvrin, Raschke & Dilla (2014), students were designated as division controllers and were required to develop an IDV to support business decisions. The case also required students to learn how to use Tableau software with minimal instructions to create the IDV.

Previous research shows that most data visualization courses cater to multi-disciplinary students (Owen *et al.* 2013; Ritsos & Roberts 2014; Sacha, Stoffel, Stoffel, Kwon, Ellis & Keim 2014). For instance, Kohlhammer's data visualization course at the Technical University of Darmstadt consisted of students with majors in computer science, mathematics, engineering, business informatics, and psychology (Owen *et al.* 2013). Programming was a significant part of this curriculum and required students to engage in real-world practical exercises. The course consisted of both technical and non-technical students, although the course content had a more technical perspective. In contrast, Owen *et al.* (2013) dicuss a course by Ebert and Elmqvist on Introduction to Visual Analytics that targeted a multidisciplinary student population and required no programming knowledge. Rather, students were required to have knowledge of one of the following: data analysis, knowledge management, statistics, computer graphics, or visualization. The course was less technical and more research-focused. Learning outcomes resulted in research presentations at conferences.

Previous studies have noted the skills gap in data visualization competencies and have attempted to design relevant curriculum to meet this need (Ritsos & Roberts 2014; Nolan & Perrett 2016; Lo, Ming & Qu 2019). For most applications, the role of data visualization requires that a curriculum not only utilizes the often-multidisciplinary skills possessed by the students, but also teaches multidisciplinary content. Data visualization is no longer only an art; it requires scientific techniques and skills as well. For instance, scientific concepts such as the Gestalt Principle of Perception and Visual Attention are central to the design of an effective data visualization output (Yang, Dai & Zhang 2022).

Research on how to design and evaluate an effective data visualization course or programme to close the skills gap in data visualization is lacking. To this end, this paper contributes to data visualization pedagogy by designing and evaluating a data visualization course. The target population is multidisciplinary learners with different backgrounds such as accounting, finance, and management information systems.

## CURRICULUM DESIGN

### Learning Objectives of the Course

In developing the learning objectives, consideration was given to the target group of students with a mixture of technical and non-technical backgrounds. The students with technical backgrounds had prior experience with analytics-based courses. They were familiar with data management tasks such as data acquisition, cleaning, and variable reduction. Students with non-technical backgrounds were comfortable with generating domain knowledge about data, but they lacked data management skills.

Four overarching goals were formulated with specific learning objectives. The four overarching goals were:

- **Overview and justification of data visualization:** where students were expected to appreciate the background and role of data visualization in decision-making.
- **Data management capabilities:** where students could acquire, clean, and organize data.
- **Principles, techniques, and best practices:** where students could build appropriate static charts and interactive data visuals.
- **Communication:** where students could communicate generated insights efficiently.

The specific learning objectives and their respective overarching goals are categorized in Table 1 below.

***Table 1:*** *Learning Objectives and Respective Categories*

| Learning Objective | Category |
| --- | --- |
| Students should be able to explain an overview and history of the practice of data visualization | Overview and justification |
| Students should be able to justify the need for visualization as a tool for decision-making. | Overview and justification |
| Students should be able to develop an ability to obtain, clean, and prepare data for visualization | Data management |
| Students should be able to develop skills to analyze complex data sets | Data management |
| Students should be able to explain concepts and principles of data visualization | Principles, techniques, and best practices |
| Students should be able to develop skills to design and create visualizations using software tools | Principles, techniques, and best practices |
| Students should be able to communicate the outcome of visual analysis in a way that would support decision-making | Communication |

**Course Structure**

Table 2 shows the course structure that informed the design of the course syllabus. The course was designed for a 16-week schedule with two concurrent tracks. These modules can be adapted and re-organized to suit the length of any curriculum or course syllabus. Track 1 started with the introduction and sensitization of data visualization and subsequently delved into theories and principles of visualization. These were mainly executed in the form of lectures, student presentations, and discussions. Track 2 covered hands-on labs where students gained hands-on experience by using the Tableau visualization software to build visual applications.

For some modules, the topics covered under both tracks 1 and 2 were related, where theories and principles were covered under track 1 and hands-on practice was completed under track 2. For instance, under chart types in modules 4 and 5, both basic and advanced chart types were covered. We covered chart types such as bar charts, histograms, choropleth maps, geographic maps, and dot plots. We also discussed how charts types can be used to encode different data types. This discussion further covered the strengths and weaknesses of each chart type. Further, specific examples to indicate how and why different chart types would be suitable or unsuitable for representing different types of data were provided. The students finally used the Tableau visualization software to generate different charts using a business data set.

***Table 2:*** *Course Structure*

| Module | Track 1 (Theories and Principles) | Track 2 (Hands-on Labs) |
|---|---|---|
| 1 | Introduction and background<br>Exploratory versus explanatory data visualization | Introduction to Tableau software<br>Importing data into Tableau |
| 2 | Building stories with visualization<br>-Creating a narrative<br>-Understanding your audience | Data types: dimensions and measures |
| 3 | Injecting context into data story telling | Data aggregation, creating joins and relationships |
| 4 | Chart types – Part 1 | Creating charts |
| 5 | Chart types – Part 2 | Creating charts |
| 6 | Preliminary project review | |
| 7 | Visual perceptions | Creating filters |
| 8 | Midterm exam | |
| 9 | Visualization best practices | Creating filters |
| 10 | Building blocks, design principles and pre-attentive attributes | Calculations in Tableau |
| 11 | Use of color in data visualization<br>Color vision deficiency | Shapes, colors, and images in Tableau |
| 12 | Typography in data visualization | Leveraging fonts in Tableau<br>Labeling and presentation |
| 13 | Interactive visualization (dashboards) and infographics | Introduction to dashboards<br>Publishing charts |
| 14 | Using maps to represent data | Creating maps using location-based data |
| 15 | Final project presentation | |
| 16 | Final exam | |

With regard to topics related to visual creation, the two main concepts of data visualization - IDV and SDV – were split into five individual techniques as is shown in Table 3 (Ndukwe & Daniel 2020).

***Table 3:*** *Visualization Techniques Used*

| Visualization Technique | Examples |
|---|---|
| Statistical graphs | Bar chart, pie chart |
| Table visualization | Cross tab, matrix |
| Time-based visualization | Line graph, Gantt chart, real-time chart |
| Network visualization | Social network graph, tree map |
| Spatial visualization | Geographic heat map |

**Pedagogic Approach of the Course**

Based on past literature, a three-pronged pedagogic approach was used in the teaching and learning process (Asamoah *et al.* 2017). The components of the three-pronged approach are sensitization, theory/principles, and practice. The sensitization component occurred much earlier in the course, concurrently with the other two components. Students were introduced to the history and background of data visualization. Whereas visualization has existed as an art since earlier civilizations (Török & Török 2019), it was imperative that students understood both its artistic and scientific evolution over multiple generations, and as an important decision-making technique.

Next, underlining concepts in data management and visuals creation were presented. With respect to visual creation, emphasis was placed on best practices for using different data visualization techniques for specific outcomes. The work of leading visualization scholars, such as Edward Tufte (Tufte 2001), was also incorporated in the examination of how to combine text, graphics, and tables as mechanisms of presenting information.

The practical component was dominated by hands-on experience, working with real-life datasets, and generating relevant insights based on defined business objectives. A key sub-goal of the course was the communication of results and insights to learners. This sub-component is vital to analytics-based implementations, where effective communication of insights enable decision makers to design effective strategies (Alhadad 2018). Another key principle emphasized was the ability to build a seamless narrative based on the analysis. This is termed as “storytelling”(Wang, Dingwall & Bach 2019).

**Evaluation**

Learners’ understanding and competence were assessed using four main deliverables: a midterm exam, a final exam, weekly labs, and a course project. The midterm and final exams were non-comprehensive. The labs covered specific topics under discussion in each particular week. The project was comprehensive and required learners to apply concepts discussed during the entire semester. The structure for the weekly labs and project is presented in subsequent sections.

*Weekly Labs*

The weekly labs served as an opportunity for students to apply the course concepts in a piecemeal fashion. Each lab required students to use the Tableau visualization software to access and explore business data. They were submitted individually, although students were permitted to discuss their ideas with their colleagues. A sample of the weekly labs is provided in Appendix B. Below is a summary of the key elements for the labs:

- *Description*: For each lab, learners applied the concepts discussed in a particular week.
- *Focus*: The foci of the learning objectives were data acquisition, data cleaning and formatting, charts, and dashboard creation using visualization best practices.
- *Expected output*: The lab was accompanied with specific questions about a set of data. The output required students to develop both static and interactive visualizations. They were also required to prepare a short memo to some questions. For instance, they may be required to generate a suitable chart to represent a variable that was coded with geographic data. They are then queried on why they would use a specific chart type to represent the given data (for example, students may need to explain why they have chosen a geographic map over a heat map).

- *Data*: Students were given specific data sets in an excel format. In most cases, they were required to consolidate multiple data sets, clean the data, create derived variables, and make other changes depending on the task at hand.

*Course Project*

The project was an opportunity for the students to apply derived theoretical knowledge and practical experience to a real-life problem. Students were encouraged to form multidisciplinary teams themselves. This way, the students were able to take advantage of their colleagues' technical and non-technical skills in analyzing the problems and finding suitable solutions. Creating collaborative groups with a multidisciplinary element, especially for a visualization task, has been shown to foster creative diversity and innovation (Domik 2011).

Different projects were designed for students for each offering of the course. The project description of one of the course offerings is provided in Appendix A. A summary of the key elements for all projects is provided below:

- *Description*: Students used publicly available COVID-19 data as they explored the nature, spread, and impact of the COVID-19 pandemic since Spring 2020. The students defined a geographic boundary at the onset of the project where they focused on data from multiple countries, or only certain regions within a specific country. The groups also identified a specific target for the impact of their results. For instance, one group could be interested in the impact of COVID-19 on sports activities within a certain region, while another group could focus on COVID-19's impact on tourism.
- *Focus*: The foci of the learning objectives were data acquisition, data cleaning and formatting, charts, and dashboard creation using visualization best practices.
- *Expected output*: Students were expected to develop both static and interactive visualizations. They were also required to prepare a memo that explained their project's objective, data sets used, relevant questions asked of the data, and insights generated.
- *Data*: COVID-19 related data was used. Students were provided with preliminary data sources to start with. Depending on the domain a group chose (e.g., tourism) they could explore more relevant data from other sources to support their analyses.

For all projects, while the instructor provided sample sources of initial data sets to students for their analyses, students could seek other supporting datasets to bolster their analyses. For instance, some groups elected to explore the effect of COVID-19 on tourism in certain geographic regions in the United States. Therefore, the group sought tourism-related data from the target locations and examined how the spread of COVID-19 cases and related deaths hampered tourism activities in those locations.

## METHOD: ASSESSMENT OF LEARNING OBJECTIVES AND EXPERIENCES

Using an Internet-based survey instrument, students' perception of the pedagogical design for teaching data visualization was assessed to determine if the design improved their perceived data visualization competency. The instrument was adapted from Asamoah, Doran & Schiller (2020), and Brownell, Hekmat-Scafe, Singla, Chandler Seawell, Conklin Imam, Eddy, Stearns & Cyert (2015). A convenience sampling method was used for data collection. The course was open to students from all disciplines, although enrollment mostly consisted of business school students. A sample population of 121 students who were enrolled in the course was targeted. The survey instrument as shown in Table 4 consisted of 12 close-ended questions. The pre-test instrument consisted of two main categories. Items in category one comprised demographic questions while items in category two assessed the course's learning objectives using a Likert scale. The post-test utilized a Likert scale as well and comprised 27 questions in four categories. Categories one and two consisted of the post-test version of the items in the pre-test. Categories three and four, shown in Table 5, focused on students' learning experiences, and respectively zoomed into the course's components and general structure. The estimation time to complete both surveys was 20 minutes.

***Table 4:*** *Categories and Items for both Pre-test and Pos-test Items*

| Category | Item number | Item |
|---|---|---|
| Demographics (category 1) | 1 | What is your gender?<br>a) Male b) Female |
| | 2 | What is your age?<br>a) under 20 b) 20-24 c) 25-29 d) 30-34 e) 35-39 f)40-44 g) 45-49 h) 50 and over |
| | 3 | What primary degree are you pursuing?<br>a) Undergraduate b) Graduate |
| | 4 | For undergraduate students, what is your class level?<br>a) Freshman b) Sophomore c) Junior d) Senior |
| | 5 | For undergraduate students, what is your major?<br>a) Bachelors in Accountancy b) Bachelors in Economics c) Bachelors in Entrepreneurship d) Bachelors in Finance e) Management f) Bachelors in Management Information Systems g) Bachelor in Marketing h) Bachelors in Supply Chain Management |
| | 6 | For graduate students, what is your major?<br>a) Masters in Accountancy b) MBA c) Information Systems d) Logistics and Supply Chain e) Marketing Analytics and Insights f) Social and Applied Economics |
| Learning objectives (category 2) | | On a scale from 1(low) to 5 (high) to what extent are the following true: |
| | 7 | I am proficient in using Tableau for visualization |
| | 8 | I can process data (clean, format, join etc.) thoroughly for visualization |
| | 9 | I understand the background and history of visualization |
| | 10 | I understand the role of visualization to support decision-making in my discipline |
| | 11 | I understand how to choose different visualization techniques/chart types to represent a specific data and problem |
| | 12 | I am comfortable in preparing and communicating visualization results in a way that supports decision-making |

Specifically, category three assessed students' experience with individual course components using items adapted from Brownell et al. (2015). The items were categorized into different themes: *course project, multidisciplinary collaboration and communication, theory, principles, concepts, exam*, and *weekly practice*. Assessing the effect of multidisciplinary collaboration on the learning experience was essential as the target population consisted of students from multiple backgrounds and majors. Also, it was deemed worthwhile to assess students' learning experience with the communication component of the course. A key soft skill for an efficient data visualization professional is the ability to effectively communicate results and insights through avenues such as reports and presentations. To that end, whether students were satisfied with the curriculum designed to improve their communications skills was also assessed. In category four, using items adapted from Asamoah, Doran, & Schiller (2020), students' general impression about the course's structure such as workload and testing procedures was also assessed. A 5-point Likert-type scale ranging from low (1) to high (5) was used in the instrument.

***Table 5:*** *Categories and Items for Post-test Items Only*

| Category | Item number | Item | Theme |
|---|---|---|---|
| Learning experience - Course components (category 3) | | On a scale from 1(low) to 5 (high), to what extent did the following aspects of the course improve your knowledge about data visualization? | |
| | 13 | Collecting and managing data as part of the course project | Course project |
| | 14 | Analyzing and creating charts as part of the course project | |
| | 15 | Communicating the results of the course project in a class presentation | Multidisciplinary collaboration and communication |
| | 16 | Group collaborations to complete class assignments and course project | |
| | 17 | Weekly theory lectures and presentations | Theory, principles, and concepts |
| | 18 | Student presentations | |
| | 19 | Exam | Exam |
| | 20 | Weekly hands-on labs | Weekly practice |
| Learning experience - General structure (category 4) | | On a scale from 1(low) to 5 (high) to what extent are the following true: | |
| | 21 | I learned a lot in this course | General learning experience |
| | 22 | The workload was appropriate for the hours of credit | |
| | 23 | Assignments were relevant and useful | |
| | 24 | Testing and evaluation procedures were good | |
| | 25 | Students were adequately involved | |
| | 26 | This course was worthwhile to me | |
| | 27 | Overall, this was a good course | |

## ANALYSES AND RESULTS

In accordance with the research aim, objectives, and research question(s), this study sought to assess the design and development of a curriculum to improve data visualization skills. The analyses and the results are presented below.

**Statistical Analyses and Results**

In total, 67 learners responded to the survey. Table 6 presents a summary of the responses to the demographic questions. The table indicates that most of the respondents were undergraduate seniors. All the graduate students were Accountancy majors, while the undergraduate students were mostly Accountancy, Finance, Management Information Systems, and Marketing majors. The others were Economics, Management, or Supply Chain Management majors. Most of the learners were males, while a majority ranged between 20 and 24 years of age.

***Table 6:*** *Summarized Demographic Table*

| Item | Percentage | Item | Percentage |
|---|---|---|---|
| **Gender** | | **Degree** | |
| Male | 60.29 | Undergraduate | 82.35 |
| Female | 39.71 | Graduate | 17.65 |
| **Age** | | **Undergraduate major** | |
| Under 20 years | 0.00 | Accountancy | 17.86 |
| 20-24 years | 79.41 | Economics | 5.36 |
| 25-29 years | 13.24 | Entrepreneurship | 0.00 |
| 30-34 years | 2.94 | Finance | 17.86 |
| 35-39 years | 4.41 | Management | 5.36 |
| 40-44 years | 0.00 | Management Information Systems | 35.71 |
| 45-50 years | 0.00 | Marketing | 14.29 |
| 50 years and over | 0.00 | Supply Chain Management | 3.57 |
| **Graduate students' major** | | **Undergraduate degree level** | |
| Accountancy | 100.00 | Freshman | 0.00 |
| Information Systems | 0.00 | Sophomore | 0.00 |
| Logistics and Supply Chain | 0.00 | Junior | 7.14 |
| Marketing Analytics and Insights | 0.00 | Senior | 92.86 |
| Social and Applied Economics | 0.00 | | |

Table 7 shows the results of the assessment of students' learning objectives. Students' perceived competencies in data visualization concepts and practice before and after the course were examined. Using paired t-tests, the study examined if students' perceived competencies improved at the end of the course after exposing them to the course's pedagogical content and design. As shown in Table 7, each learning objective has a corresponding question item in the survey except for the objectives under the data management category. This was because students' competencies in data acquisition, cleaning and preparation encompasses complex data sets as well, and hence one question item was sufficient in testing the two objectives under the data management category. Table 7 indicates that for all items, students perceived significant improvements in their competencies ($p<0.0001$).

***Table 7:*** *Pre-Test and Post-Test Mean Comparisons for Learning Objectives*

| Item number | Category | Item | Pre-mean | Post-mean | p-value |
|---|---|---|---|---|---|
| 7 | Overview and justification | I understand the background and history of visualization | 1.91 | 3.72 | <0.0001 |
| 8 | Overview and justification | I understand the role of visualization to support decision-making in my discipline | 3.00 | 4.30 | <0.0001 |
| 9 | Data management | I can process data (clean, format, join etc.) thoroughly for visualization | 2.28 | 3.75 | <0.0001 |
| 10 | Principles, techniques, and best practices | I understand how to choose different visualization techniques/chart types to represent a specific data and problem | 2.67 | 4.25 | <0.0001 |
| 11 | Principles, techniques, and best practices | I am proficient in using Tableau for visualization | 1.55 | 3.57 | <0.0001 |
| 12 | Communication | I am comfortable in preparing and communicating visualization results in a way that supports decision-making | 2.45 | 4.00 | <0.0001 |

In addition to the learning objectives, Table 8 portrays students' learning experiences for specific components of the course. These items were only in the post-test. The table shows the mean scores and standard deviations of the responses. Table 8 also reflects the percentage of responses that were rated as 4 and 5. Except for the exam component, the mean scores of all the responses exceed 3.5. Also, for all the components except the exam, most of the respondents rated the component as either a 4 or a 5. While the respondents favorably rated the exam component, they did not find it an overwhelmingly relevant course component.

***Table 8:*** *Learning Experience Results About Specific Course Components*

| Item number | Item | Mean | Std. dev. | 4 and 5 ratings (%) |
|---|---|---|---|---|
| 13 | Collecting and managing data as part of the course project | 3.76 | 1.05 | 62.12 |
| 14 | Analyzing and creating charts as part of the course project | 4.15 | 0.91 | 72.72 |
| 15 | Communicating the results of the course project in a class presentation | 3.85 | 1.03 | 62.12 |
| 16 | Group collaborations to complete class assignments and course project | 3.92 | 1.09 | 66.70 |
| 17 | Weekly theory lectures and presentations | 3.53 | 1.26 | 53.08 |
| 18 | Student presentations | 3.71 | 0.95 | 59.09 |
| 19 | Exam | 3.17 | 1.05 | 33.34 |
| 20 | Weekly hands-on labs | 3.92 | 1.09 | 72.73 |

***Table 9:*** *Results for General Course Learning Experience*

| Item number | Item | Mean | Std. dev. | 4 and 5 ratings (%) |
|---|---|---|---|---|
| 21 | I learned a lot in this course | 3.95 | 0.91 | 69.70 |
| 22 | The workload was appropriate for the hours of credit | 4.17 | 0.88 | 78.78 |
| 23 | Assignments were relevant and useful | 4.02 | 0.84 | 74.24 |
| 24 | Testing and evaluation procedures were good | 3.70 | 1.06 | 60.61 |
| 25 | Students were adequately involved | 3.91 | 0.95 | 66.67 |
| 26 | This course was worthwhile to me | 4.12 | 0.96 | 74.24 |
| 27 | Overall, this was a good course | 4.20 | 0.87 | 80.30 |

With respect to students' general learning experiences, Table 9 shows a mean score that is approximately 4. In addition, except for testing and evaluation procedures, close to 70% of the respondents rated each item in the top 2 categories. As compared to the other learning experiences, a smaller number of respondents favored the testing and evaluation procedures.

## DISCUSSION

In accordance with the research aim, objectives, and research question(s), this study sought to assess the design and development of a curriculum to improve data visualization skills. The study also explored key components for effectively teaching a data visualization course. Previous studies suggest that data visualization have been mainly taught in computer science curricula due to the computational skills needed to build modern charts (Nolan & Temple Lang 2010; Ceri 2018; Ryan *et al.* 2019). Industry has made similar assumptions and implications, given the overwhelming amount of job postings that list data visualization skills in tandem with traditional computer science skills such as programming and web development (Ryan *et al.* 2019). However, this narrative has broadened to include other disciplines (including business).

Given the inherent requirement for both technical and non-technical skills for generating deep insights about data, benefits of data visualization as an analytics tool can be accessed by multiple disciplines, including business. Knowledge gained could be rewarding to a mix of both technical and non-technical learners, given that teams who accomplish data visualization tasks in organizations require diverse and overlapping skills (Walny, Frisson, West, Kosminsky, Knudsen, Carpendale & Willett 2019). As shown in Table 6, there was general interest from multiple disciplines in the course. Factors such as accreditation requirements and required skills for accountancy, finance, management information systems, and marketing jobs have increased students' interest in learning more about the application of data visualization to decision-making.

As a research objective, the study explored if the curriculum design and pedagogic approach improved students' perceived competence in data visualization skills. The research objective of this study correlates with the learning objectives identified for the course as shown in Table 1. Furthermore, the learning objectives were operationalized with specific survey questions as shown in Table 7.

The results show that the curriculum design helped students achieve competency in data processing and preparation, which is a key requirement for effective data analysis. In addition, students were able to leverage requisite domain knowledge to effectively prepare data for analysis.

Consistent with the data analytics discipline, encouraging students to understand and use the requisite domain knowledge was vital to generating relevant visuals during data analysis (Waller & Fawcett 2013). The course design also helped students improve their proficiency in visualization techniques to create relevant charts to support decision-making. Students understood when and why to choose different types of charts and visualization techniques to achieve a specific outcome. This decision-making skill is vital for generating and preparing effective reports and communicating relevant results to stakeholders so they can make the right business and technical decisions.

The data in Table 6 shows that, whereas a majority of students were management information systems majors, there was an appreciable number of students from other disciplines such as finance and accountancy. The interdisciplinary structure of the course made it feasible for students to benefit from the data visualization concepts taught in the course. Students from multiple backgrounds had the opportunity to interreact and share ideas on relevant assignments which is consistent with what some of the prior research studies have promoted for a data visualization curriculum (Elmqvist & Ebert 2012; Owen *et al.* 2013). Interdisciplinarity is vital since traditionally, data visualization concepts borrow from a disparate and broad spectrum of courses. Elmqvist & Ebert (2012) noted that an interdisciplinary approach to data visualization should leverage topics in areas such as knowledge discovery, cognitive and perceptual science, statistical analysis, and data management. These concepts were discussed as part of the topics outlined in the course. Learners agreed that their skills in the related topics improved significantly.

In Table 7 the data shows that learners significantly improved on their skills in using specialized Tableau visualization software to generate relevant visualization products. This is consistent with data visualization needs in both academia and industry (Ryan *et al.* 2019; Walny *et al.* 2019). The emphasis on Tableau hands-on competency was included in the course because, whereas underlying theories and concepts may be persistent as compared to ever-changing software tools, the software tools presented a viable means to operationalize the theories discussed in the course. Besides, since Tableau is one of the leading data visualization software in the data analytics discipline, students who gain relevant software knowledge can hit the ground running and assimilate easily into new roles that require data visualization skills. For instance, in a survey conducted by Ryan *et al.* (2019), Tableau was listed as a job requirement in 41% of job postings out of a total of 30,000 job postings that mentioned data visualization.

In addition to understanding if students' competencies significantly improved based on the curriculum design, students' perception about their learning experiences in the course was also explored. These learning experiences were categorized into two main parts. First, students learning experiences regarding the individual components of the course, and next, students' general learning experiences about the course. The survey instrument was adapted from Asamoah, Doran & Schiller (2020) and Brownell *et al.* (2015).

Students perceived the various components of the course as vital to favorably enhancing their learning experiences (Table 8). For instance, data collection and management, data analyses and visual creation, and communication of results were some of the most highly favored components of the course. This is a significant outcome because these three components constitute some of the key functions of a data analytics project (Azevedo & Santos 2008). Consequently, skills learned in the data visualization course can be transferred by learners to improve an organization's entire data analytics posture and investment.

In addition, exercises and group collaboration were highly-favored components of the course. Students favored the hands-on and practical nature of the curriculum design. The ability to gain not just theoretical knowledge, but also practical skills is an integral component of several visualization and analytics-based courses (Mcquaigue, Burlinson, Subramanian, Saule & Payton 2018; Qasim & Kharbat 2020). Learners found the hands-on structure of the weekly labs to be critical since it helped them to easily apply relevant concepts to the course project. This is consistent with prior research which shows that by including hands-on and group collaborative assignments in a curriculum, learners are typically able to assimilate concepts much faster and longer (Taylor & Rohrer 2010; Luo 2020; Maeko 2020).

Relative to other components, however, students neither favored the exam component nor the testing procedures in the course as shown in Tables 8 and 9 respectively. This could be attributed to the expectation that students generally do not perceive exams as an effective evaluation method (Sletten 2021). Exams mostly focus on memorization of visualization knowledge whereas other assessment methods such as labs, enhances learners' ability to analyze, evaluate, and generate solutions to problems. Also, as compared to other forms of assessment such as course projects and term paper reviews, exams do not allow flexibility and easy transition of courses to alternative delivery modes in emergency situations (Sletten 2021). An example is the sudden pivot to virtual course delivery by academic institutions during the COVID-19 pandemic. Hence, the use of exams as a key evaluation component for data visualization courses should be carefully reviewed by instructors. In situations where its use is imperative to the overall delivery of the course content, instructors may consider assigning less percentage points to it as part of the overall course grade.

Generally, the students agreed that they learned a lot in the course and that the workload and assignments were respectively appropriate and relevant for their learning (Table 9). They perceived their involvement, especially in the group projects, as vital to improving their data visualization competence. Finally, they also perceived the assignments as relevant and useful.

## CONCLUSION

In summary, data visualization is common in academia and in research environments. However, its application has been shown to be vital in industry applications and for managerial decision-making as well (Heer & Shneiderman 2012). Previous studies have shown that data analytics and business intelligence constitute some of the key Information Technology (IT) investment areas in organizations (Thouin, Hefley & Raghunathan 2018; Wang, Yeoh, Richards, Wong & Chang 2019; Qasim & Kharbat 2020).

In this study, a curriculum was designed for teaching data visualization as a pillar for data analytics and business intelligence in organizations and businesses. The pedagogic approach using a survey instrument was also tested. The results show that leaners perceived an increased competency in data visualization skills due to the data visualization curriculum approach and design. Students also generally favored the various components of the course as well as the structure. The learners agreed that their learning experiences were enhanced.

Although data visualization courses are not new in academia, there have been several forms proposed without support for how such courses will improve leaners' skills as far as the application of data visualization concepts to data analytics is concerned (Mcquaigue *et al.* 2018). This study is significant because it presents a data visualization course that was designed and further tested to ensure that learners' skills would be improved. This study also helps chart the path for future pedagogical studies in data visualization. It will help both academia and industry practitioners decide what relevant components to include in a data visualization course.

In recognizing the limitations of this study, it is noted that while data visualization is a fundamental skill for analyzing data and generating insights, there are other related skills that enhance its effectiveness. Skills in data management and programming, such as Structured Query Language and Python, are relevant to data visualization applications as well. However, the role of such related skills and tools were not accounted for in this study. Subsequent studies should focus on a broader view of data visualization pedagogy, with respect to accompanying tools and skills that enhance its effectiveness.

Also, as a limitation, the curriculum was tested on a convenience sample of students. Even though this could affect the external validity of the study, as a pedagogical study, a student sample was relevant for the study’s objective. When the results are extended to the development of data visualization curriculum in a different context, a relevant and more representative sample should be used so as to maintain the external validity of the results.

In both academia and industry, curriculum in data visualization is being promoted as a relevant analytics tool for leveraging IT capabilities for decision-making (Kokina, Pachamanova & Corbett 2017). Data visualization, as part of business analytics is a resource-intensive tool (Wang *et al.* 2019). In order for an organization to maximize its IT investments in data visualization, employees with the right skills should be trained to generate innovate insights from business data. This course supports such efforts by designing a data visualization curriculum that would help improve the data visualization skills of leaners from multiple backgrounds.

## REFERENCES

Adams, A., Boersema, T. and Mijksenaar, M. (2010). "Warning symbology". *Information Design Journal*, vol. 18, no. 2, pp. 94–106. doi: 10.1075/idj.18.2.01ada.

Alhadad, S. S. J. (2018). "Visualizing data to support judgement, inference, and decision making in learning analytics: Insights from cognitive psychology and visualization science". *Journal of Learning Analytics*, vol. 5, no. 2, pp. 60–85.

Asamoah, D. A., Sharda, R., Hassan Zadeh, A. and Kalgotra, P. (2017). "Preparing a data scientist: A pedagogic experience in designing a big data analytics course". *Decision Sciences Journal of Innovative Education*, vol. 15, no. 2, pp. 161–190. doi: 10.1111/dsji.12125.

Asamoah, D. A., Doran, D. and Schiller, S. (2020). "Interdisciplinarity in Data Science Pedagogy: A Foundational Design". *Journal of Computer Information Systems*, vol. 60, no. 4, pp. 370–377. doi: 10.1080/08874417.2018.1496803.

Azevedo, A. and Santos, M. F. (2008). "KDD, SEMMA and CRISP-DM: A Parallel Overview" in *IADIS European Conference Data Mining*, pp. 182–185. Available at: http://recipp.ipp.pt/handle/10400.22/136.

Brownell, S. E., Hekmat-Scafe, D. S., Singla, V., Chandler Seawell, P., Conklin Imam, J. F., Eddy, S. L., Stearns, T., Cyert, M. S. (2015). "A High-Enrollment Course-Based Undergraduate Research Experience Improves Student Conceptions of Scientific Thinking and Ability to Interpret Data". *CBE-Life Sciences Education*, vol. 14, no. 2.

Ceri, S. (2018). "On the role of statistics in the era of big data: A computer science perspective". *Statistics & Probability Letters*, vol. 136, pp. 68–72.

Domik, G. (2011). "Fostering Collaboration and Self-Motivated Learning: Best Practices in a One-Semester Visualization Course". *IEEE Computer Graphics and Applications*, vol. 32, no. 1, pp. 87–91.

Dong, T. and Triche, J. (2020). "A Longitudinal Analysis of Job Skills for Entry-Level Data Analysts". *Journal of Information Systems Education*, vol. 31, no. 4, p. 312.

Elmqvist, N. and Ebert, D. S. (2012). "Leveraging multidisciplinarity in a visual analytics graduate course". *IEEE computer graphics and applications*, vol. 32, no. 3, pp. 84–87.

Erbacher, R. F. and Forcht, K. A. (2010). "Combining visualization and interaction for scalable detection of anomalies in network data". *Journal of Computer Information Systems*, vol. 50, no. 4, pp. 117–126.

Gupta, B., Goul, M. and Dinter, B. (2015). "Business Intelligence and Big Data in Higher Education: Status of a Multi-Year Model Curriculum Development Effort for Business School Undergraduates, MS Graduates, and MBAs". *Communications of the Association for Information Systems*, vol. 36, no. 23, pp. 449–476.

Heer, J. and Shneiderman, B. (2012). "Interactive Dynamics for Visual Analysis". *Queue*, vol. 10, no. 2, pp. 30–55.

Janvrin, D. J., Raschke, R. L. and Dilla, W. N. (2014). "Making Sense of Complex Data using Interactive Data Visualization". *Journal of Accounting Education*, vol. 32, no. 4, pp. 31–48.

Kokina, J., Pachamanova, D. and Corbett, A. (2017). "The Role of Data Visualization and Analytics in Performance Management: Guiding Entrepreneurial Growth Decisions". *Journal of Accounting Education*, vol. 38, pp. 50–62.

Lo, L. Y. H., Ming, Y. and Qu, H. (2019). "Learning Vis Tools: Teaching Data Visualization Tutorials", in *2019 IEEE Visualization Conference, VIS 2019*. Vancouver: IEEE. doi: 10.1109/VISUAL.2019.8933751.

Luo, B. (2020). "The influence of teaching learning techniques on students' long-term learning behavior". *Computer Assisted Language Learning*, vol. 33, no. 4, pp. 388–412.

Maeko, M. S. A. (2020). "Teaching and Learning Through Hands-on Activities in Civil Technology Teacher Training Programmes at South African Universities". *Journal of Human Ecology*, vol. 71, no. 1–3, pp. 38–48. doi: 10.31901/24566608.2020/71.1-3.3182.

Mcquaigue, M., Burlinson, D., Subramanian, K., Saule, E. and Payton, J. (2018). "Visualization, assessment and analytics in data structures learning modules", in *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*. Baltimore, pp. 864–869.

Miah, S. J., Solomonides, I. and Gammack, J. G. (2020). "A design-based research approach for developing data-focussed business curricula". *Education and Information Technologies*, vol. 25, no. 1, pp. 553–581.

Microsoft (2020). "Power BI". Redmond, WA: Microsoft.

Näykki, P. and Järvelä, S. (2008). "How pictorial knowledge representations mediate collaborative knowledge construction in groups". *Journal of Research on Technology in Education*, vol. 40, no. 3, pp. 359–387. doi: 10.1080/15391523.2008.10782512.

Ndukwe, I. G. and Daniel, B. K. (2020). "Teaching Analytics, Value and Tools for Teacher Data Literacy: A Systematic and Tripartite Approach". *International Journal of Educational Technology in Higher Education*, vol. 17, no. 1, pp. 1–31.

Nguyen, A., Gardner, L. and Sheridan, D. (2020). "Data analytics in higher education: an integrated view". *Journal of Information Systems Education*, vol. 31, no. 1, p. 61.

Nolan, D. and Perrett, J. (2016). "Teaching and Learning Data Visualization: Ideas and Assignments". *The American Statistician*, vol. 70, no. 3, pp. 260–269.

Nolan, D. and Temple Lang, D. (2010). "Computing in the statistics curricula". *The American Statistician*, vol. 64, no. 2, pp. 97–107.

Odewumi, O. M. (2021). "Empowering Students' Cognitive Learning of Creative Colours through Computer-Based Concept Maps". *International Journal of Education & Development using Information & Communication Technology*, vol. 17, no. 3, pp. 155–165.

Owen, G. S., Domik, G., Ebert, D. S., Kohlhammer, J., Rushmeier, H., Santos, B. S. and Weiskopf, D. (2013). "How Visualization Courses Have Changed Over the Past 10 Years". *IEEE Computer Graphics and Applications*, vol. 33, no. 4, pp. 14–19.

Panjwani, S., Micallef, L., Fenech, K. and Toyama, K. (2009). "Effects of integrating digital visual materials with textbook scans in the classroom". *International Journal of Education and Development using ICT*, vol. 5, no. 3, pp. 55–71.

Qasim, A. and Kharbat, F. F. (2020). "Blockchain technology, business data analytics, and artificial intelligence: Use in the accounting profession and ideas for inclusion into the accounting curriculum". *Journal of emerging technologies in accounting*, vol. 17, no. 1, pp. 107–117.

Ritsos, P. D. and Roberts, J. C. (2014). "Towards more visual analytics in learning analytics", in *Proceedings of the 5th EuroVis workshop on visual analytics*. Swansea, pp. 61–65.

Ryan, L., Silver, D., Laramee, R. S. and Ebert, D. (2019). "Teaching Data Visualization as a Skill". *IEEE Computer Graphics and Applications*, vol. 39, no 2, pp. 95–103.

Sacha, D., Stoffel, A., Stoffel, F., Kwon, B. C., Ellis, G. and Keim, D. A. (2014). "Knowledge generation model for visual analytics". *IEEE transactions on visualization and computer graphics*, vol. 20, no. 12, pp. 1604–1613.

Sletten, S. R. (2021). "Rethinking Assessment: Replacing Traditional Exams with Paper Reviews". *Journal of Microbiology & Biology Education*, vol. 22, no. 2. doi: 10.1128/jmbe.00109-21.

Srinivasa, R. C., Arun, S., James, L. and Zhu, Y. (2021). "Visualization - a vital decision driving tool for enterprises". *arXiv preprint arXiv:2015.11037*. Available at: http://arxiv.org/abs/2105.11037 (Accessed: 21 October 2021).

Stenliden, L., Bodén, U. and Nissen, J. (2019). "Students as Producers of Interactive Data Visualizations—Digitally Skilled to Make Their Voices Heard". *Journal of Research on Technology in Education*, vol. 51, no. 2, pp. 101–117. doi: 10.1080/15391523.2018.1564636.

Stodder, D. (2013). *Data visualization and discovery for better business decisions*, *TDWI Research*. Available at: http://www.solutiondesignteam.com/wp-content/uploads/data-visualization-discovery-better-business-decisions-106672.pdf.

Tableau (2020). "Tableau Software". Seattle, WA: Tableau.

Taylor, K. and Rohrer, D. (2010). "The effects of interleaved practice". *Applied Cognitive Psychology*, vol. 24, no. 6, pp. 837–848.

Thouin, M. F., Hefley, W. E. and Raghunathan, S. (2018). "Student attitudes toward information systems graduate program design and delivery". *Journal of Information Systems Education*, vol. 29, no. 1, p. 25.

Török, Z. G. and Török, Á. (2019). "Cognitive Data Visualization—A New Field with a Long History", in Klempous, R., Nikodem, J., and Baranyi, P. Z. (eds) *Cognitive Infocommunications, Theory and Applications*. Springer, Cham, pp. 49–77. doi: 10.1007/978-3-319-95996-2_3.

Tufte, E. R. (2001). *The visual display of quantitative information*. Graphics Press. doi: 10.1109/MPER.1988.587534.

Waller, M. A. and Fawcett, S. E. (2013). "Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management". *Journal of Business Logistics*, vol. 34, no. 2, pp. 77–84.

Walny, J., Frisson, C., West, M., Kosminsky, D., Knudsen, S., Carpendale, S. and Willett, W. (2019). "Data Changes Everything: Challenges and Opportunities in Data Visualization Design Handoff". *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 1, pp. 12–22. doi: 10.1109/TVCG.2019.2934538.

Wang, S., Yeoh, W., Richards, G., Wong, S. F. and Chang, Y. (2019). "Harnessing business analytics value through organizational absorptive capacity". *Information & Management*, vol. 56, no. 7.

Wang, Z., Dingwall, H. and Bach, B. (2019). "Teaching data visualization and storytelling with data comic workshops", in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, pp. 1–9.

Ward, M. O., Grinstein, G. and Keim, D. (2010). *Interactive Data Visualization: Foundations, Techniques, and Applications*. 1st edn. New York: CRC Press.

Wixom, B., Ariyachandra, T., Goul, M., Gray, P., Kulkarni, U. and Phillips-Wren, G. (2011). "The Current State of Business Intelligence in Academia". *Communications of the Association for information Systems*, vol. 29, no. 1, pp. 299–312.

Yang, Z., Dai, Q. and Zhang, J. (2022). 'Visual perception driven collage synthesis". *Computational Visual Media*, vol. 8, no. 1, pp. 79–91.

Copyright for articles published in this journal is retained by the authors, with first publication rights granted to the journal. By virtue of their appearance in this open access journal, articles are free to use with proper attribution, in educational and other non-commercial settings

## APPENDIX A: SAMPLE COURSE PROJECT

### Project Title: Analyzing COVID-19 Using Data Visualization

#### *Introduction*

COVID-19 has altered how we generally live and do business. Many organizations have had to shut down or only partially operate due to the impact of this virus. At the high level, the purpose of this case is to creatively analyze relevant COVID-19 data and present insights that would help understand the nature, spread, and impact of the disease.

#### *Tasks and Strategies*

To proceed with your analysis, consider the following as a starting point. Be creative and reach beyond the ideas suggested in this case description.

- Your analysis should help identify, explore, explain, and predict a phenomenon related to COVID-19.
- It is your responsibility to zoom in and ask specific questions related to COVID-19. The questions can further be refined once you start reviewing the data sources and analyzing your data.
- Your ability to ask the right questions, locate the right data, clean the data, perform appropriate visualization techniques, and interpret the results is vital in this exercise.
- You want to provide key insights that individuals, organizations (e.g. schools) and governments can rely on as they draft best-practices, guidelines, and policies to guide their operations.
- For instance, based on trends you show, should governments enact blanket shut down policies across countries, states, cities etc. If not, will a staggered approach to shut down policies be more effective. Are there any thresholds for disease outbreaks that should be targeted? Can you show how comorbidities factor into infections and deaths?

To help perform efficient analyses, consider the strategies below:

- Ask the right questions.
    - You can start with broad level questions, but it is important to narrow down to a few basic questions or topics that will provide added insights. To do this, you can narrow down your analysis based on number of countries, time etc. That is, you can analyze several variables over a short period, or you can analyze a few variables over a long period of time. The question generation stage can be iterative.
- Extract, transform and load relevant data.
    - Domain knowledge is important in understanding the data. Also, some spend time in understanding the variables in the data.
- Apply appropriate data analytics (visualization) techniques.
    - Think about what type of analysis and charts will be appropriate for the story you intend to tell your chosen audience.
- Interpret and share the results with stakeholders.
    - Keep the visual simple, clean, and easily interpretable. Sometimes, less is more. Do not clutter the dashboard. Show one change at a time in animations. The audience should easily understand what question is being addressed from your visualization.

***Data***

Data can be extracted from a plethora of sources. Below, I provide three key sources although you are welcome to include data from other sources to enhance your analysis.

- http://www.healthdata.org/covid/data-downloads
- https://github.com/owid/covid-19-data/tree/master/public/data
- https://covid.cdc.gov/covid-data-tracker/?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fcoronavirus%2F2019-ncov%2Fcases-updates%2Fcases-in-us.html#learn-more

Data from these websites are related to country, continent, population, age, new cases, deaths, tests, hospitalizations, and infections. Information on trends, forecasts and demographics are also provided. Beyond raw data, reports are also provided in some cases. You are at will to use data from just one or all sources provided.

***Expected Output Rubric***

Your deliverable should consist of a combination of visualization and a memo. The visualization should be expressed in a dashboard, infographic, animation, or a combination of these. The memo should be a formal report that highlights the key findings and observations generated with the visualizations. Table A1 shows a rubric for assessing the visualization output.

***Table A1:*** *Rubric for Assessing Data Visualization Output*

| Grading Category | Excellent | Good | Unsatisfactory |
|---|---|---|---|
| Basic requirements (15%) | At least 3 unique charts/maps are used in the visuals. Are dashboards included. | Overly complex or simple visuals. Some relevant charts/maps are missing. | The visuals are significantly longer or shorter than necessary. Does not use multiple unique charts. |
| Ideas and insight (25%) | The visuals answer clear and relevant specific questions. | The visuals answer clear specific questions that are somewhat relevant. | The visuals answer questions that are unclear and irrelevant. |
| Visuals (35%) | The visuals clearly illustrate key insights from the dataset and follow data visualization principles and best practices. | The visuals illustrate some insights from the dataset and follow most data visualization principles and best practices. | The visuals do not illustrate key insights from the dataset and/or do not follow data visualization principles and best practices |
| Anticipating questions and interactivity (25%) | The visuals that permit filtering or highlighting and allows user interactions based on anticipated relevant questions. The most interesting findings are highlighted using preset filters. | The visuals that permit filtering or highlighting and allows limited user interactions based on anticipated relevant questions. | The visuals that do not permit filtering or highlighting, and/or does not allow user interactions based on anticipated relevant questions. |

## APPENDIX B: SAMPLE OF WEEKLY LAB

In this lab, you would use concepts in "Data types: dimensions and measures" to answer the following questions. Where applicable, create one visualization sheet for each task.

- **Task 1**
  Connect to Inventory.xlsx data file. How many entity types are there in the data set?
- **Task 2**
  Name 1 each of an ordinal, ratio and interval variables in the data set.
- **Task 3**
  Which entity type accounted for the most sales? Create a calculated field to sum sales (label the calculation "Total_Sales").
- **Task 4**
  Using a filter on sales, which city had the highest sales?
- **Task 5**
  How much baseball hats (in dollars) was sold in the city with the most sales?
- **Task 6**
  Using the same filter, which state sold the most items? Show your results on a geographic map?
- **Task 7**
  Explain how you used a filter to answer Task 4 and 6.
- **Task 8**
  What entity type was returned the most in the Other category? What was the dollar value?
- **Task 9**
  What entity type accounted for the least sales in the Other category?
- **Task 10**
  What kinds of entities are included in the Miscellaneous category? Explain how you arrived at your conclusion.

## APPENDIX C: SOFTWARE RESOURCES FOR DATA VISUALIZATION

***Table C1:*** *Software and Tools for Data Visualization*

| Software | Source | Notes |
|---|---|---|
| Tableau | https://www.tableau.com/academic | Academic use available |
| SAS Visual Analytics | https://www.sas.com/en_us/learn/academic-programs.html | Academic use available |
| Microsoft Power BI | https://powerbi.microsoft.com/en-us/ | Free trial available |
| Qlik Sense | https://www.qlik.com/us/products/qlik-sense | Academic use available |
| Microstrategy | https://www.microstrategy.com/en | Academic use available |
| R statistical software | https://www.r-project.org/ | Open source |
| Python | https://www.python.org/ | Open source |
| Gephi | https://gephi.org/ | Social Network Analysis<br>Link Analysis<br>Academic use available |
| NodeXL | https://www.smrfoundation.org/nodexl/ | Social Network Analysis<br>Academic use available |