INFORMATION SYSTEMS EDUCATION JOURNAL

Volume 21, No. 5
November 2023
ISSN: 1545-679X

In this issue:

Issue 5 has six contributions to the field focusing in many cases on less-served populations, and seeking paths toward improved services and access. This includes mobile banking in Ghana, teaching cases on US rural health care and on design thinking to improve access to community services, and assessing online vs. on-ground IS education. Our final two papers renew our focus on IS education fundamentals, with attention to tool choices for spreadsheet education, and student perception of online tutors.

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 EDSIGCON (https://proc.iscap.info) 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 36%.

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 members of ISCAP who perform the editorial and review processes for ISEDJ.

INFORMATION SYSTEMS EDUCATION JOURNAL

Editors

Exploring Methods Cybersecurity Managers Need to Implement to Minimize Cyber-Frauds in Mobile Money Services in Ghana

Bright Siaw Afriyie
bsiawa@sabsoftech.org
Sab-Softech Inc
1129 Beechwood Ln, Cedar Hill, TX 75104

Samuel Sambasivam
Samuel.Sambasivam@Woodbury.edu
Computer Science Data Analytics
Woodbury University
Burbank, CA 91504

Abstract

Nearly half the adult population in developing nations lack a formal bank account and other financial services. Ghana is no exception, having a massive community of unbanked adults and among those countries positioned at the bottom of the spectrum of financial inclusion. The unbanked populations unfortunately had to use physical cash-carry in every business transaction. The advent of mobile financial services (MFS) provides an alternative mode to financial inclusion. MFS, fundamentally implemented using SMS and unstructured supplementary service data (USSD) code, essentially encompasses mobile wallets, cash out, and over-the-counter transactions overgrowing globally, has a vast potential to minimize impediments to financial inclusion. However, MFS' sustainability is threatened by cybercriminals or fraudsters while reaching increasingly a higher global penetration yearly. This qualitative exploratory study explores the top security methods cybersecurity managers need to implement to minimize the cyber fraud of mobile financial services in Ghana. The research purports to identify viable methods leaders of MFS operators need to implement to reduce fraudsters' threats. The exploratory design was used as the lens to explore this phenomenon in-depth. A sample size of seven interviewees and 12 semi-structured interview questions were used as a data collection instrument. Lack of proper security methods and internal control processes were identified as the major causes of cyber fraud in MFS. Seven databases (Web of Science, ProQuest, ABI, EBSCO, IEEE, Sage, Google Scholar, Pub/Med, and Scopus) were searched using standard and adapted search syntax. The study also provided four recommendations on the ecosystem and the processes needed to utilize mobile technology's full potential.

Keywords: Mobile Network and Financial Service, Mobile Money Fraud, Mobile Security Architecture, Unbanked Financial Inclusion, Cryptography.

1. INTRODUCTION

The global economy is increasingly expanding mobile financial services. Developing countries represent the global leader in mobile financial services (Akomea-Frimpong et al., 2019). Mobile money popularly referred to as "MoMo" or "mobile wallet," is a financial technology (FinTech) offering unbanked people the possibility to use a mobile phone to transfer, receive, deposit, and spend

money rather than carrying physical cash. Depending on the country, mobile money is named based on specific services such as mPesa, EcoCash, GCash, Tigo Pesa, etc. According to Nicco-Annan (2020), there were over 270 mobile money services worldwide; however, they were most present in Africa, Asia, and Latin America. In developing nations, where the unbanked population was significantly high, MoMo emerged as a suitable alternative to cash, formal banking, and financial transactions since it was considered easy to use, secure and accessible anywhere there was a mobile phone signal (Akomea-Frimpong et al., 2019; Nicco-Annan, 2020). The mobile financial service covers a full spectrum of financial services, from payment transactions and checking to savings accounts, loans, insurance, and investments (Chironga et al., 2017). The proliferation is expected to extend to media and entertainment, web services, retail, and healthcare by 2024 (Ingle, 2019; Normans Media Ltd, 2017).

The MFS has predominantly gained roots in developing nations, particularly in Africa. The MFS operations in the developing countries is uniquely different from Venmo, PayPal, CashApp, etc., since these processes are tied to banking accounts. MFS relies uniquely on SMS transmission using mobile phones without any ties with banking accounts. In a sub-Sahara country like Ghana, the number of cellular or mobile registered voice subscribers in March 2020 stood at 41,959,298 (National Communications Authority, 2020). Out of 32.7 million registered mobile money accounts, 14.7 million remain active, with 235,000 active operator agents (Bank of Ghana, 2020) for all four major mobile service providers operating in the country. According to Boateng (2018), about 83.1% of Ghanaians have mobile money accounts, while Naomi Kwetey et al. (n.d.) asserted that the unbanked population was estimated at 70% of the entire population. Mobile Telecommunications Network (MTN), a leading company and the first to introduce MFS in 2009, now has 23,945,672 registered subscribers, representing 57.07% of registered subscribers in Ghana. Vodafone followed in 2011 with 8,787,464 subscribers, representing 20.94% of registered subscribers in Ghana, Airtel-Tigo with 8,498,008 subscribers representing 20.25%, and Glo with 728,154 subscribers representing 1.74% of registered subscribers (National Communications Authority, 2020).

With about 13.05 million active subscribers, the mobile money industry realized a cash transaction value estimated at GHC223.2bn (US$43bn) in 2018. In 2019, this value jumped to GHC265.42bn, with 25.78 million active registered subscribers being 26% over the yrevious year (Botchey et al., 2020). The drive continues as mobile money's market size expects to grow from USD 21.15 billion in 2016 to USD112.29 billion by 2021, and MoMo payments are expected to compound an annual growth rate (CAGR) of 22% by 2024 (PR Newswire, 2019). IndustryARC (2019) projected a CAGR of over 24.9% growth in the forecast period 2021 to 2026.

The Ghana Interbank Payment and Settlement Systems (GhIPSS) reported that mobile money interoperability (MMI) increased by 358% during the first quarter of 2020 (Nicco-Annan, 2020). This upsurge is causing an exponential expansion of the cyber surface, creating numerous opportunities for cyber fraudsters (Weber & Studer, 2016). The cyber fraudsters continued to exploit vulnerabilities associated with deficiencies in proper security in the mobile money market; this has immensely attracted the need to identify and develop appropriate countermeasures against cyber fraud threats (Singh & Kumar, 2018). In Sub-Saharan Africa, including Ghana, 564 million mobile phone users sustained a cellular penetration rate of 65% in 2013. The figure was expected to reach 947 million in 2020, a penetration rate of 91%. Akomea-Frimpong et al. (2019) posited that mobile money services transactions lacked internal controls, efficient methods, and tools to curb the threats. Kanobe et al. (2017) affirmed the existence of literature gaps regarding MFS's information security management in developing economies, particularly reviews on the mobile money industry insider employees' roles in securely protecting the MFS business. According to Kyeremeh (2018), Delta3 International reported that in 2016 Ghana lost 50.0 million USD to cyber-related attacks. In Africa, the figure was approximately 2.0 billion USD –(GhanaWeb).

The study used the exploratory qualitative method to identify effective security methods cybersecurity managers need to implement to minimize digital fraud in mobile financial services in Ghana (Guma et al., 2020). In essence, it intended to provide recommendations to increase the safety of vulnerable populations engaging in electronic financial transfers in developing countries at risk. Akomea-Frimpong et al. (2019) articulated that MFS fraud is a salient economic problem; however, little has been captured in the research literature. The study further sought to fill this literature gap. The research used Game Theory to explain the phenomenon's underlying concepts. The theory binds the concept based on

the Diffusion of innovations and the model of planned behavior (Weigel et al., 2014) to relate to Cybersecurity and information systems assurance research.

2. HISTORICAL FACTS OF THE UNBANKED

The historically unbanked population had suffered in a way to deliver profitable business, as they had to carry physical cash. They banked their funds in their homes and carried cash for all business remittances and acceptance. Global remittances have been slow, less effective, inefficient, and very expensive for consumers, yet remittances are crucial for so many businesses worldwide (Bettman & Harris, 2014). Small and medium-sized businesses typically unbanked have been challenged with armed robbery and burglary in many instances while transporting physical discharge of their routine business activities. MFS represents a digital disruption that has started to move the industry forward through the opportunity of mobile devices to drive revolutionary change in the US$550bn remittance market, driving exponential change. With MFS, there is no more inherited infrastructure and expensive processes in doing business. Described as an innovative FinTech, MFS is reshaping the digital market structure, how investors and consumers receive and use the information and financial services, and how companies access and deploy capital. MFS processes ranging from new digital payment systems and digital or electronic currencies to online investment and finance platforms and data analytics are already impacting traditional financial markets and services (Brummer & Gorfine, 2014).

3. RESEARCH PROBLEM

The problem is the lack of understanding and availability of efficient and effective methods that cybersecurity managers need to implement to minimize cyber fraud in mobile money services in Ghana (Akomea-Frimpong et al., 2019; Guma et al., 2020; Novelan et al., 2018; Singh & Kumar, 2018). The lack of proper protection methods or MFS not adequately protected creates opportunities for cyber fraudsters to hack mobile systems and steal customers' money. Bank of Ghana reported that Ghana lost over GHc 12.8 million (USD 2.2 million) in 2021 to mobile money-related fraud (Thebftonline.com,2022). Security methods remain a salient problem in mobile money systems regarding using short message services (SMS) and unstructured supplementary service data (USSD) with inherent vulnerabilities (Novelan et al., 2018). Eavesdropping and intercepting money transfer data in transit posed significant security concerns for several mobile money ecosystems. No scheme or existing method could offer complete SMS and SS7 security (Khozooyi et al., 2009; Novelan et al., 2018). Lee and Narayanan (2021) found that out of 259 samples of recycled mobile phone numbers available to new subscribers in the United States at two significant carriers, 171 were tied to existing accounts, exposing security and privacy risks to potential hackers to exploit. No substantial research has been conducted to reduce these threats. The menace continued to seriously affect mobile transactions' integrity (Maseno et al., 2017). Kisekka (2019) noted that according to MTN Uganda, some suspicious individuals obtained PINs from customers under pretenses and subsequently withdrew funds from customers' mobile money accounts.

The pertinence of the problems articulated led to the research question: *Q1*: What methods do cybersecurity managers need to implement to minimize cyber fraud in mobile money services in Ghana?

4. RESEARCH METHODOLOGY

This qualitative exploratory study explores top security methods cybersecurity managers must implement to minimize cyber fraud in Ghana's mobile financial services (MFS) (Akomea-Frimpong et al., 2019; Guma et al., 2020). Qualitative exploration was best suited as it attracted securing in-depth information-rich data (Bazen et al., 2021; Leedy & Ormrod, 2016). Adopting the Constructivist approach (McSweeney, 2018) focused on the target sample population of lived experience. Exploratory Design was used as a lens to capture the in-depth information stemming from 12 semi-structured interview questions from seven participants (Levitt, 2021). The exploratory Design satisfied the choice of interview participants in this qualitative study to produce additional meaning that made more sense.

5. CONCEPTUAL FRAMEWORK

The underlying conceptual framework for this study was based on Game Theory. Game theorists strive to understand conflicts and cooperation in more complicated real-life situations to comprehend real competitive conditions (Smith, 1982). According to Osborne (2004), game theory endeavors to help understand situations in which opposing decision-makers interact. Generally, a game is a competitive activity in which players compete based on rules. Myerson (2013) described game theory's language as a game

referring to any social situation which involves two or more individuals called players. In-game approach, the game theorists highlight two fundamental assumptions that players are made up of (a) rational and (b) intelligence. The gamers also represent rational decision-makers if they make decisions in the consistent pursuance of their objectives. Building on judgment, the theory's fundamental results assume that each player aims to maximize the expected value of the payoff measured on a specific utility scale. This decision, according to Amoroso & Magnier-Watanabe (2012), identified eleven key-related variables which have been blended into nine: (1) perceived ease of gaming and risk; (2) attitude toward game; (3) facilitating conditions; (4) perceived value and usefulness; (5) perceived security and privacy; (6) social influence; (7) trust; (8) behavioral intention to game; (9) the attractiveness of alternatives. The idea is that a rational decision- maker must make decisions based on maximizing the expected utility payoff (Von Neumann & Morgenstern, 1947) via favorable winning methods.

Based on Kolokoltsov & Malafeyev's (2020) affirmation, every human being strives to accomplish their essential goals at any stage while establishing permanent contact with others trying to achieve their purposes. People sometimes attempt to outsmart an opponent and occasionally establish alliances with cronies with similar interests. The theory was constructed under the premise of committing no mistakes. An atom of a single error or negligence of any player would certainly favor its opponent.

Similarly, bringing the concept down to a cybersecurity protection environment, flaws in any defensive mechanisms would favor the attacker who consistently searched for them. The framework illustrated that humans remain in a strict perpetual dichotomy in cyber warfare in which everyone strives to survive by winning over the other party. Game Theory concept outlines three parties struggling in cyberspace: defenders or security experts, cybercriminals or intruders, and the target game zone comprising society, industries, government, and academia, which must be protected against cyberattacks.

6. POPULATION AND SAMPLE

The target population or universe comprised a set of people of interest that provided appropriate responses to the research question (Naseri, 2021; Sekaran & Bougie, 2010). The target population more suited for this study were people in mobile network operations with lived experience relating to the phenomenon under investigation (Creswell, 2014; Osborne & Grant-Smith, 2021). The target demographics group was at least 18 years of age with a minimum of three years of work experience in cybersecurity or mobile network operations. The participatory population included professional cybersecurity managers and mobile network operator (MNO) managers in mobile phone corporations. Seven out of 10 purposeful samples recruited for this qualitative exploratory design were more suited for in-depth interviews. Getting a group targeting or sampling from a broader population range in various organizations provided diversity; primarily, selecting participants through the random sample addressed trustworthiness. Additionally, the purposeful sampling used for most qualitative research could address research bias and eliminate any unknown influences in the group targeting process (Editorial Board, 2016, p.47).

7. DATA ANALYSIS

To provide a meaningful understanding, an in-depth analysis of the data produced two categories, principal themes and popular topics, which summarized the critical points from the data collected. The principal themes represented the essential matters of the subject participants mainly elaborated on their insight. Popular topics covered less frequently mentioned findings but remained the main topics aligned with and strived to answer the central research question: what methods do cybersecurity managers need to implement to minimize cyber fraud in mobile financial services in Ghana?

Principal Themes and Popular Topics	50	100%
Participants Major Themes		
Security Improvement Methods	18	36%
Vulnerability Detecting	12	24%
Risk and Threat Assessment	7	14%
On-Boarding Process	5	10%
Phishing Scam Control	4	8%
Popular Topic		
Operations Improvement Process	4	8%

Table 1: Participants Data Principal Themes and Popular Topics

The six themes signified the main point of this study and interpreted a clear sense of the aggregated data collected as indicated in table 1. From the impartial assessment, six major

themes emerged from the data, from which one popular topic was identified. These themes provided an understanding of the phenomenon, and isolated a central topic focusing on the methods for operational improvements.

The themes were generated from Nvivo 12 Pro software with codes principally centered on cybersecurity methods and mobile security. The principal themes represented the key elements that were highlighted by the subject participants, based on their personal insights.

Principal Theme 1: Security Improvement Methods

The study unraveled that the mobile money ecosystem consisted of three principal actors ascertaining what Guo and Bouwman (2016) pointed out: the mobile users or holders of money accounts, the service provider or telecommunication companies, and the mobile money agents (Botchey et al., 2020). The methods used to detect mobile money fraud have been based on rules. The security improvement methods predominantly touched the operations of these three stakeholders in the MFS business. As the focus was on improving security in the mobile systems to be as resilient as possible, it would entail implementing at least baseline controls, including access controls, risk management, and personal security (Mtaho & Mselle, 2014). A baseline security process centered on the CIA triad configuration, embracing confidentiality, integrity, and availability (Shin et al., 2013) would require more stringent access authentication configurations such as two-factor authentication (2FA) and secure password mechanisms. The study has revealed that swapped SIM would work in any cellular phone registered in the same carrier or any unlocked phone from a different carrier. Still, unless the swapper obtains the PIN Code from the original SIM's registered owner, the attacker cannot access the victim's mobile account. A visible interaction aspect of the security objectives on IT business objectives was the frequent patches updates, which often slowed down system performance. Patches often installed late in schedules became less effective.

The study found that, due to technology evolution and compatibility requirements, almost all mobile telecoms in Ghana implement the GSM basic architecture blended in 2G, 3G, and 4G or LTE. It is worth pinpointing that none of the telecoms in the nation deploys a 5G network. While telecoms in Ghana embrace 2G and 3G technologies, advanced countries like the USA have AT&T telecom company discontinued deployment of 3G technology that could no longer support VoLTE (Voice over LTE) as of February 2022, according to Spectrum TV news channel. These telecom firms view the profitability only on the business side, embracing low-end devices without considering the security vulnerabilities such technologies embed. Mobile Networks using outdated technologies harbor vulnerabilities that attract many security threats like DDOS (Yin et al., 2018), brute force attacks, and system intrusions, with unsolicited SMS text messages and fake promotions ads. Security improvements could be implemented by making most of the 5G architecture features embedding software-defined network (SDN) technology (Yao et al., 2019).

The mobile network runs Advanced Encryption Standard (AES) encryption algorithms with a key length greater than 128-bits. However, among several attack vectors against AES encryption, brute force appears as a known attack (Burr, 2003), and biclique cryptanalysis (Bogdanov et al., 2011) can lead to breaching the protection of mobile services. As Bogdanov et al. (2011) asserted and confirmed in a recent Sowmiya and Malarvizhi (2022) study about security for machine intelligence-based cryptosystems. Using an increasingly AES encryption key length of 256 bits would guarantee acceptable and trusted standards in the mobile systems would improve security. AES-256 appeared to offer one of the most difficult challenges in block cipher cryptanalysis for over a decade, requiring $2^{254.4}$ computational complexity to break it.

Principal Theme 2: Vulnerability Detecting

The methods for detecting mobile money fraud have been based on rules (Yadav & Sora, 2021). As revealed in this study, fraud detection also leverages the top 10 OWASP and CVSSv3 vulnerabilities daily guides. The tools often used were penetration testing, threat modeling, intrusion detection and prevention systems (IDS)/(IPS), endpoint detection devices, and firewall filtering systems. The massive mobile money transactions (MMTs) also presented potential security challenges to telecom firms. Machine and deep learning methods have lately emerged as efficient financial tools to detect and prevent fraudulent activities (de Sá et al., 2018; Hajek & Henriques, 2017; Sadgali et al., 2019; Singh et al., 2012). The support vector machine (SVM), k-nearest neighbor (knn), and artificial neural network (ANN) are machine learning algorithms experts could deploy to predict possible fraud occurrences and prevention (Saxena et al., 2019).

Fraud detection thus plays a vital role in providing firm data assurance. Ponnusamy et al. (2020) enounced that social engineering's catastrophic effect has caused the industry to build adequate security infrastructure. The mobile network operations (MNO) employees must receive sufficient awareness training to combat social engineering attacks (Salahdine & Kaabouch, 2019). It was mind-boggling to discover that most MFS agent operators in Ghana did not receive formal security awareness training Adopting Big Data Analytics (BDA) methods, as Gupta et al. (2018) noted, has the potential to detect fraud. Detecting critical fraud data, BDA allows verifying fraud insights to react in real time, investigate suspicious activities and analyze historical data based on fraud and financial crime patterns (Ranjan & Foropon, 2021). The study supported the highlights pronounced by Guma et al. (2020) regarding digital money mobile account security being prone to severe vulnerabilities affecting the safety of MoMo.

The USSD holds severe vulnerabilities. However, it functions as a real-time protocol for mobile communication technology used in GSM network architecture to deliver supplementary services for text exchange among mobile handset devices and application programs, avoiding access to the internet (Nyamtiga et al., 2013; Talom & Tengeh, 2019). The SS7 supports signaling to network elements without a direct trunk connection (Lei et al., 2021; Paganini, 2016). Discovery in this study was associated with participants' assertion that the SS7 protocol has a long-standing vulnerability that has never been fixed. In GSM architecture, weak encryption of broken stream cipher is optional, which occurs only with the traffic in the airway between the Mobile Station (MS) and the BTS or the cellular tower. The message contents spread in the air in plaintext at transmission time without any intrinsic security countermeasure (Chaeikar et al., 2021). Undeniably, this process presents a serious risk because the providers' operation staff could access and manipulate the SMS information on the Short Message Service Center (SMSC). Therefore, SMS usage for mobile money transfers has enormous security concerns (Chaeikar et al., 2021), exposing threats such as eavesdropping, interception, and text modification.

Principal Theme 3: Risk and Threat Assessment

The study findings confirmed that although the MNOs engaged in an inhouse Risk Assessment Framework for detecting and mitigating vulnerabilities effectively, they leaned heavily toward using the NIST SP 800-30 and ISO27005 Risk Assessment framework as a central focus of their security program. The findings revealed that by applying the threat models described in NIST SP 800 53 standards, mobile network operating firms would be better prepared to detect and mitigate threats (NIST, 2020). Besides, threat identification was based on knowledge or skills, risk assessment, and utilization of threat modeling. The threat modeling appeared to pull tremendous knowledge from empirical research findings on attackers' activities all over the globe. The industry conducts risk assessments every quarter. One of the best practices in sourcing data from threat Intelligence would be to focus on potential threats, evaluate the patterns, and learn how these threats could occur, especially in reconnaissance.

Principal Theme 4: On-Boarding Process

Emulating the advanced countries' successes, Ghana launched a national identification registration scheme as the backbone of the technology initiative for her citizens' security (Stranek-Africa, 2019). The national identification card (NIC) has become the preferred identity card for mobile phone registration. The study discovered that the telecom industry's onboarding process covers mobile phone SIM card registration methods and leverages a nationwide campaign for mobile subscribers to re-register their SIM cards using the NIC as the valid form of identity card. The implication is that the NIC has additional biometric data embedded, strengthening the security features for the onboarding process. Figure 1, an illustrative overview diagram, provides a simplified interpretation of the onboarding process for mobile money services.

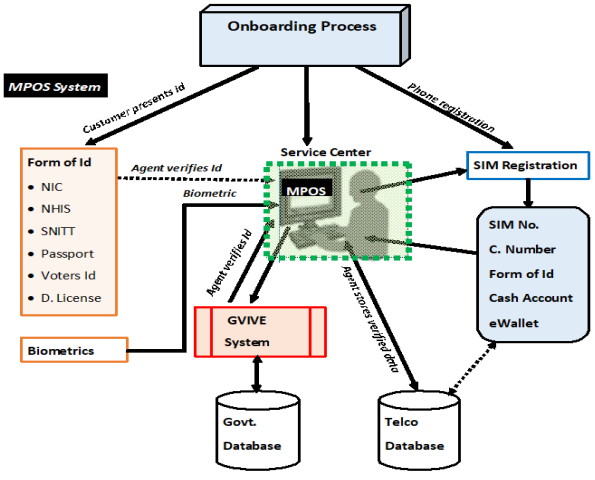


Figure 1: The Onboarding Process Diagram Derived from the Participants' Responses

Registration must reference a SIM Card and Mobile Number and information on a valid Government Issued Identity Card including demographic data like the legal name, date of birth (DOB), Identity Card number, and the next kin or heir. In situations where the credentials do not match, the telecom's certified registration manager will authenticate the process by asking security questions to verify the authenticity of the Identity card presented by the subscriber. When the customer and the agent agree to load the money, the initial MoMo sending process begins right after onboarding. All data is stored on the telecom company's data servers as the background information for the mobile money account.

Principal Theme 5: Phishing Scam Control

Phishing scams stood up as the most severe mobile money services threat ravaging every facet of the MFS business in this study. The menace appeared as a high-tech scam that manifested itself in different ways of fraudulent actions, including deceptive calls, mobile channel extortion, internet extortion, mobile fraud, identity theft, or hacking. The study showed that phishing incidents usually occurred when the scammers tricked subscribers into disclosing their secret PIN codes through a phone call, SMS text messages, or embedded hidden malware in an email link (Asha & KR, 2021). Additionally, fraudsters discovered a means to steal users' credit card sensitive data when they send fake mobile SMS or place calls through impersonation. As identified in this study, the fraud occurred when fraudsters employed social engineering tools like cited phishing tactics to trick mobile subscribers and agents when fraudsters pretended to use valid registered mobile numbers to call their victims (Wrigley, 2018).

Other forms of phishing identified in this study were smishing and vishing attacks when fraudsters used an emotional delusional SMS message to trick users into revealing their mobile money PIN (Maseno et al., 2017). Once executed, the fraudster exploits the existing flaws and sends an SMS to trick the user (victim) into confirming a payment when no money has been transferred.

8. RECOMMENDATIONS

The study deduced that the world's economy is increasingly expanding mobile financial services through online and mobile platforms. Consumers deem MFS convenient at any location with a wireless signal. Past research has revealed scant literature on a holistic approach to reducing mobile cash fraud. The rise in cashless transactions through mobile devices is hitting skyrocketing levels in everyday business transactions critical to developing economies. There has been no thorough method to secure mobile financial services to minimize security exposure and prevent fraud. These threats have caused some telecom (providers) and their subscribers to lose billions of dollars. USSD and SMS as a mode of transmission for financial data coupled with SS7 protocol present enormous security concerns. GSM architecture has inherent encryption algorithm flaws, and MFS operations lacking adequate internal controls have led to the four recommendations to reduce these threats.

Recommendation 1

The practical implications of this study focus on methods and processes for fraud prevention. The holistic techniques in this study's contribution centered on adequately securing the PIN Code to Mobile Accounts and inadequate method of implementation of security defense systems from a practical perspective. Factors affecting the security of the PIN Code depend on three actors mentioned earlier in this document, including cybersecurity experts, agent operators, and subscribers. These factors reflect the practical sense of Game Theory with the innovation adoption-behavior model, where the findings of this study clearly articulate the concept of the stringent dichotomy of perpetual cyberwarfare existing between cybersecurity experts' world and cyber fraudsters. The cybersecurity experts must perform their duties in establishing the appropriate security methods mandating strict policies and procedures in mobile network operations. Cybersecurity managers must develop policies involving top managers, agent operators, and subscribers regarding the use of their mobile network. Engaging the subscribers in policymaking could be accomplished through open surveys and customers' feedback.

Recommendation 2

The implications remained on the need to mandate biometric-based identity cards like national identity cards for onboarding and cash withdrawal processes. The risk assessment method embedding threat modeling was a valuable tool for detecting most vulnerabilities and potential threats. A great discovery has shown that the daily risk assessments conducted within the mobile network infrastructure have been quite helpful. The most critical finding was using encryption technology to protect users' privacy. The study revealed AES encryption algorithm with a key length greater than 128-bits was used. The precise key length of at least 256

bits is recommended as such key length can provide adequate security. Scheduled patches' installation must be respected regardless of business profitability impediments. Owing to the evolution of technology in the face of fraudsters' sophistication advancement, cybersecurity managers must retire the use of mobile architectural generation lower than 4G/LTE. Instead, they must deploy a 5G network blended with SDN architecture (Niyaz et al., 2017).

Recommendation 3

The mobile agent operators who are savvy in mobile technology must not take advantage of their customers. Cybersecurity experts must restrict agent operators and affiliated contractors from the administrator access to mobile systems. Restricting agent operators from installing privileges can prevent replay attacks by installing software to withdraw smaller amounts from mobile subscribers' accounts. Prevent manipulating the SMS text by making it read-only when conveying money transfers to subscribers' accounts and prevent extorting customers' mobile money accounts. Agent operators must receive adequate security awareness training to combat phishing scams and to allow agent operators to defend themselves and the entire telecom industry. Management must devise a standard form of identity verification, especially in a cash withdrawal transaction. Strict measures must cover any cash transfers mandating agent operators to collect cash from the sender customer before executing and committing the money transfer process.

Recommendation 4

The subscribers must comply with the "You For Know" program designed to guide customers to protect the privacy of their PIN Code to their mobile money accounts. There was also a discovery that the awareness programs developed by the mobile carriers do not reach even the agent operators, not to mention the subscribers. The subscribers must note that the only security tool for protecting their mobile accounts is their PIN Code, activating the security features on their mobile handsets, and upgrading iOS and Android to mobile application security verification standards (MASVS). The cybersecurity expert could not deliver absolute security protections to their accounts except through the two parties' collaboration.

9. CONCLUSION

The focus of this qualitative exploratory study aimed to explore top security methods cybersecurity managers must implement to minimize cyber fraud of mobile financial services in developing nations, particularly in Ghana. The exploratory design was best suited because the objective was to capture in-depth information-rich data content. The response to the research question addressed the overarching problem - the lack of efficient methods that managers need to implement to minimize cyber fraud in mobile money services. The study's findings have concluded the essence of the theoretical framework describing methods to mitigate cyber fraud as the relay goes to the fraudster or the cyber expert to win the game. The implication was a failure to provide an adequate defense to mobile systems proactively. The relay, again, goes to the fraudster or the cyber expert to win the game.

10. REFERENCES

Akomea-Frimpong, I., Andoh, C., Akomea-Frimpong, A., & Dwomoh-Okudzeto, Y. (2019). Control of fraud on mobile money services in Ghana: an exploratory study. *Journal of Money Laundering Control, 22*(2),300-317. https://doi.org/10.1108/JMLC-03-2018-0023

Amoroso, D. L., & Magnier-Watanabe, R. (2012). Building a research model for mobile wallet consumer adoption: the case of mobile Suica in Japan. *Journal of Theoretical and Applied Electronic Commerce Research, 7*(1), 94-110. https://doi.org/10.4067/S0718-18762012000100008

Asha, R. B., & KR, S. K. (2021). Credit card fraud detection using artificial neural network. *Global Transitions Proceedings, 2*(1), 35-41. http://www.keaipublishing.com/en/journals / global-transitions-proceedings/

Bank of Ghana. (2020). Summary of economic and financial data. https://www.bog.gov.gh/wp-content/uploads/2020/03/Summary-of-Economic-Financial-Data-March-2020.pdf

Bazen, A., Barg, F. K., & Takeshita, J. (2021). Research techniques made simple: An introduction to qualitative research. Journal of Investigative Dermatology, 141(2), 241-247.

Bettman, J., & Harris, M. (2014). Mobile money: The impact of smartphones on the international remittance market. *Journal of Payments Strategy & Systems, 8*(3), 264-273.

Boateng, K. (2018). Ghana's progress on reaching out to the unbanked through financial inclusion. *International Journal of Management Studies, 5*(2).

Bogdanov, A., Khovratovich, D., & Rechberger, C. (2011). Biclique cryptanalysis of the full AES. Proceedings of theory and application of cryptology and information security. *International Conference on the, in ASIACRYPT 2011: Advances in Cryptography*,344-371. https://doi.org/10.1007/978-3-642-25385-0_19

Botchey, F. E., Qin, Z., & Hughes-Lartey, K. (2020). Mobile money fraud prediction—A cross-case analysis of the efficiency of support vector machines, gradient boosted decision trees and naïve Bayes algorithms. *Information,11*(383). https://doi.org/10.3390/info11080383

Brummer, C., & Gorfine, D. (2014). FinTech: Building a 21st-Century Regulator's Toolkit

Burr, W. E. (2003). Selecting the advanced encryption standard. *IEEE Security & Privacy, 99*(2),43-52. https://doi.org/10.1109/MSECP.2003.1193210

Chaeikar, S. S., Yazdanpanah, S., & Chaeikar, N. S. (2021). Secure SMS transmission based on social network messages. International Journal of Internet Technology and Secured Transactions, 11(2), 176-192.

Chironga, M., De Grandis, H., & Zouaoui, Y. (2017). Mobile financial services in Africa: Winning the battle for the customer. *McKinsey & Co. Financial Services*.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approach– 4th edition.* Sage Publications.

de Sá, A. G., Pereira, A. C., & Pappa, G. L. (2018). A customized classification algorithm for credit card fraud detection. *Engineering Application and Artificial Intelligence* 72, 21–29.

Editorial Board. (2016). *Perspectives of Qualitative Research Methods.* Words of Wisdom, LLC.

Guma, A., Dida, M. A., & Anael Elikana, S. (2020). Evaluation of key security issues associated with mobile money systems in Uganda. *Information,11*(6),309. http://dx.doi.org/10.3390/info11060309

Guo, J., & Bouwman, H. (2016). An ecosystem view on third party mobile payment providers: A case study of Alipay wallet. *Info,* 18, 56–78.

Gupta, S., Kar, A. K., Baabdullah, A., & Al-Khowaiter, W. A. A. (2018). Big Data with cognitive computing: A review for the future. *International Journal of Information Management*, 42, 78–89.

Hajek, P., & Henriques, R. (2017). Mining corporate annual reports for intelligent detection of financial statement fraud – A comparative study of machine learning methods. *Knowledge-Based System,* 128, 139–152.

IndustryARC. (2019). *Mobile money market – Forecast 2021 - 2026.* Industry Analysis Research Consulting Report Code: ITR 0035. https://www.industryarc.com/Report/15195/mobile-money-market.html

Ingle, S. (2019). Mobile money market — Statistics, growth analysis, trends, demand, competitive landscape to 2023. *Medium*

Kanobe, F., Alexander, P. M., & Bwalya, K. J. (2017). Policies, regulations, and procedures and their effects on mobile money systems in Uganda. *The Electronic Journal of Information Systems in Developing Countries, 83*(1), 1-15.

Khozooyi, N., Tahajod, M., & Khozooyi, P. (2009). Security in mobile governmental transactions. *Second International Conference on Computer and Electrical Engineering, 2*, 168-172.

Kisekka, J. I. (2019). MTN Uganda issues a statement on mobile money fraudulent withdrawals, 2019. https://www.dignited.com/45203/mtn-statement-mobile-money-fraud-withdrawals/

Kolokoltsov, V. N., & Malafeyev, O. A. (2020). Understanding game theory: Introduction to the analysis of many agent systems with competition and cooperation. (p. 3). *World Scientific*

Kyeremeh, H. (2018). Ghana's National Cybersecurity Policy and Strategy (NCSPS): Critique and Comparison with Best Practice

Lee, K., & Narayanan, A. (2021). Security and privacy risks of number recycling at mobile carriers in the United States.

Leedy, D. P., & Ormrod, J. E. (2016). *Practical Research: Planning and Design, 11e*. Published by Pearson Education, Inc.

Lei, Z., Nan, Y., Fratantonio, Y., & Bianchi, A. (2021). On the insecurity of SMS one-time password messages against local attackers in modern mobile devices. In Proceedings of the *2021 Network and Distributed System Security (NDSS) Symposium*.

Levitt, H. M. (2021). Qualitative generalization, not to the population but to the phenomenon: Reconceptualizing variation in qualitative research. Qualitative Psychology, 8(1), 95.

Maseno, E. M., Ogao, P., & Matende, S. (2017). Vishing attacks on the mobile platform in Nairobi County, Kenya*. International Journal Advanced Research Computing and Scientific. Technology,* 5, 73–77.

McSweeney, K. (2018). Motivating cybersecurity compliance in critical infrastructure industries: A grounded theory study. (10744158 Ph.D.), Capella University.

Mtaho, A. B., & Mselle, L. (2014). Securing mobile money services in Tanzania: A Case of Vodacom M-Pesa. *Int. J. Comput. Sci. Netw. Solut., 2*, 1–11.

Myerson, R. B. (2013). Game theory: Analysis of conflict. *Harvard University Press*.

Naomi Kwetey, D. B. A., Caesar, L., Appiah, D., & Collins Cobblah, M. B. A. (n.d.). Banking the unbanked in Ghana. *SBS Journal of Applied Business Research, 44.*

Naseri, R. N. N. (2021). What is a population in online shopping research? A perspective from Malaysia. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12*(4), 654-658.

National Communications Authority. (2020). *Mobile voice subscriptions from January to March 2020*. Communications Industry Statistics. https://www.nca.org.gh/industry-data-2/market-share-statistics-2/telecom-voice/Communications-Industry-Statistics-MV-Q1.pdf

Nicco-Annan, J. (2020). *That's MoMo like it: Everything you need to know about mobile money in Ghana.* WorldRemit. https://www.worldremit.com/en/stories/story/2020/05/28/mobile-money-ghana

NIST. (2020). National institute of standard and technology special publication 800-53, Revision 5. *Natl. Inst. Stand. Technol. Spec. Publ. 800-53, Rev. 5*, 492 pages (September, 2020). https://doi.org/10.6028/NIST.SP.800-53r5

Niyaz, Q., Sun, W., & Javaid, A. Y. (2017). A deep learning-based DDoS detection system in software-defined networking (SDN). *EAI Endorsed Transactions on Security and Safety, 4*(12), Article ID 153515, 2017.

Normans Media Ltd. (2017). Mobile money market global key vendors, manufacturers, suppliers, and analysis market report 2022. *M2 Presswire*.

Novelan, M. S., Husein, A. M., Harahap, M., & Aisyah, S. (2018). SMS security system on mobile devices using a tiny encryption algorithm. *In journal of physics: conference series 1007(*1), 012037). *IOP Publishing*.

Nyamtiga, B. W., Anael, S., & Loserian, L. S. (2013). Enhanced security model for mobile banking systems in Tanzania. *Intl. Jour. Tech. Enhancements and Emerging Engineering Research, 1*(4), 4–20.

Osborne, M. J. (2004). *An introduction to game theory (Vol. 3).* Oxford University Press.

Osborne, N., & Grant-Smith, D. (2021). In-depth interviewing. *Methods in Urban Analysis,* 105-125.

Paganini, P. (2016). *SS7 Protocol: How Hackers Might Find You.* InfoSec. https://resources.infosecinstitute.com/topic / ss7-protocol-how-hackers-might-find-you/

Ponnusamy, V., Selvam, L. M. P., & Rafique, K. (2020). Cybersecurity governance on social engineering awareness. *In Employing Recent Technologies for Improved Digital Governance,* 210-236. *IGI Global.*

PR Newswire. (2019). Mobile money market to expand at CAGR of 22% till 2024, rising mobile payments drive growth; says TMR. *Transparency Market Research*. https://doi.org/201902260500PR.NEWS.USP R.IO65412

Ranjan, J., & Foropon, C. (2021). Big data analytics in building the competitive intelligence of organizations. *International Journal of Information Management, 56*, 102231.

Sadgali, I., Sael, N., & Benabbou, F. (2019). Performance of machine learning techniques in the detection of financial frauds. *Procedia Computer Science*, 148, 45–54.

Salahdine, F., & Kaabouch, N. (2019). Social engineering attacks: a survey. *Future Internet, 11*(4), 89.

Saxena, S., Vyas, S., Kumar, B. S., & Gupta, S. (2019). Survey on online electronic payments security. In Proceedings of the *2019 Amity International Conference on Artificial Intelligence (AICAI)*, 746–751

Sekaran, U., & Bougie, R. (2010). *Research methods for business: A skill-building approach, (5th ed.).* John Wiley & Son

Shin, S., Yegneswaran, V., & Porras, P. (2013). AVANT-GUARD: Scalable and vigilant switch flow management in software-defined networks. Proceedings of the *2013 ACM SIGSAC conference on Computer & Communications security*, pp. 413– 424.

Singh, G., Gupta, R., Rastogi, A., Chandel, M. D., & Ahmad, R. A. (2012). Machine learning approach for detection of fraud based on SVM. *International Journal of ScientificEngineering and Technology, 1*, 192–196.

Singh, H. B., & Kumar, G. (2018). Cybercrimes: A proposed taxonomy and challenges. *Journal of Computer Networks and Communications*. https://doi.org/10.1155/2018/1798659

Smith, J. M. (1982). *Evolution and the theory of games.* Cambridge University Press.

Sowmiya, G., & Malarvizhi, S. (2022). Design of Testability Structures with Security for Machine Intelligence-Based Cryptosystem. *SRMIST: SRM Institute of Science and Technology.* Research Square. https://doi.org/10.21203/rs.3.rs-1217671/v1

Stranek-Africa. (2019). *National identification registration will be an exercise in futility – Stranek-Africa.* Fly Multimed Ghana. https://flymultimediagh.com/2019/12/09/national-identification-registration-will-be-an-exercise-in-futility-stranek-africa/

Talom, F. S. G., & Tengeh, R. K. (2019). The impact of mobile money on the financial performance of SMEs in Douala, Cameroon. *Sustainability 12*(1), 183.

Thebftonline.com. (2022). Future of MoMo industry hinges on partnerships - CEO. *Business News* of Friday, 12 August 2022. https://www.ghanaweb.com/GhanaHomePage/business/Future-of-MoMo-industry-hinges-on-partnerships-CEO-1601423

Von Neumann, J., & Morgenstern, O. (1947). Theory of games and economic behavior. (2nd rev. ed.).

Weber, S. H., & Studer, E. (2016). Cybersecurity in the internet of things: legal aspects. *Computer Law & Security Review, 32*(5), 715–728.

Weigel, F. K., Hazen, B. T., Cegielski, C. G., & Hall, D. J. (2014). Diffusion of innovations and the theory of planned behavior in information systems research: A meta-analysis. *Communications of the Association for Information Systems, 34*(1), 619–636.

Wrigley, J. (2018). *Exploring the causes and defenses of Social Engineer in Developing nations: Using Ghana as a Case Study*. Colorado Technical University Doctorate Research Paper (108288880 DCS). https://doi.org/108288880

Yadav, A. K. S., & Sora, M. (2021). Fraud detection in financial statements using text mining methods: A review. In *IOP conference series: Materials science and engineering 1020*(1), 012012. IOP Publishing.

Yao, J., Han, Z., Sohail, M., & Wang, L. (2019). A robust security architecture for SDN-based 5G networks. Future Internet, 11(4), 85.

Yin, D., Zhang, L., & Yang, K. (2018). DDoS attack detection and mitigation with Software-Defined Internet of Things Framework. *IEEE Access, 6,* 24694–24705.

Teaching Case

Creating a Clear Vision for Rural Healthcare: A Data Analysis Exercise

Christine Ladwig
cladwig@semo.edu
Department of Marketing

Taylor Webber
tweber1s@semo.edu
Department of Marketing

Dana Schwieger
dschwieger@semo.edu
Department of Management

Southeast Missouri State University
Cape Girardeau, MO 63701, USA

Hook

Data is a powerful tool for the healthcare industry to use for managing, analyzing, and reporting on critical events in the field. The analysis of broad, salient data files aids healthcare businesses in uncovering hidden patterns, market trends, and customer preferences; these details may then be used to improve the quality and delivery of care to patients in an organization's community. This case highlights the use of data analysis as a planning tool for a mid-sized rural hospital with limited resources.

Abstract

Data is a powerful tool for the healthcare industry to use for managing, analyzing, and reporting on critical events in the field. The analysis of broad, salient data files aids healthcare businesses in uncovering hidden patterns, market trends, and customer preferences; these details may then be used to improve the quality and delivery of care to patients in an organization's community. In this case, students use simple data mining procedures to investigate issues a healthcare organization faces regarding regional and national population patterns, directions for facility and service expansion, and prospective staffing changes. The exercise highlights the use of data analysis as a planning tool for a mid-sized rural hospital with limited resources and may be used in an undergraduate or graduate level management information systems or healthcare information systems course to illustrate data analysis and visualization concepts, reporting, and data driven strategy development.

Keywords: Teaching Case, Big Data Analytics, Healthcare

1. SEEING THE BIG PICTURE

Although approximately 60 million Americans live in rural areas, access to healthcare is challenging for individuals in these locations, and rural medical providers struggle financially (HHS, 2020). Nearly 50% of rural hospitals operate at a loss. They also face provider recruitment and retention challenges, revenue pressure, and patient groups with complex medical issues (HHS, 2020). Due to these factors, it is critical for small and mid-sized rural hospitals to possess a clear understanding of the needs within their patient populations.

Because their healthcare delivery has unique obstacles and issues, rural health providers are seeking methods for better analyzing their current and prospective patients' needs and requirements. A powerful tool to accomplish this objective is the use of big data, often characterized by: Volume (a large volume in many environments), Variety (wide variety of data types), and Velocity (velocity at which data is generated, collected, and processed) (Botelho & Bigelow, 2022), Veracity (accuracy in data sets), Value (relevant data), and Variability (variation in data sources). It is a common myth that the term big data just refers to a large amount of data; it is actually the ability of complex smart data to expand the understanding of the issue under examination that truly defines the term. Big data allows an organization to move beyond its immediate borders to capture the big picture of related factors, and therefore more effectively analyze and respond to an issue.

Data mining techniques are used to analyze big data and may include "quantitative analysis, database management, data visualization and intelligent computing (machine learning, pattern recognition and artificial intelligence) for the purpose of discovering patterns that are not previously known in the data" (Jafar, Anderson, & Abdullat, 2008). Data mining is being used in the healthcare environment to unearth new knowledge and trends for both clinical and administrative decision-making (Yoo et al., 2012).

Big data analysis is being used by healthcare businesses to identify patient outcomes, better organize staffing and operations, and strategize the addition of providers and services. The use of big data may also be applied to improve the diagnosis and treatment of illnesses, and lead to better prevention of disease and patient safety (Pastorino et al., 2019). Because of its breadth and source variety, big data is ideal for aiding rural healthcare institutions in addressing disparities in patient access or identifying allocation directions for scant resources. Big data collected during the height of the COVID-19 pandemic has been especially valuable in helping small and mid-sized hospitals with planning for future healthcare delivery (Kent, 2021) such as operational strategy, staffing, grant development, and service delivery.

2. THE NEED FOR CLARITY

Midwest Hospital operates as a not-for-profit 55-bed acute care facility in an area with a population of approximately 40,000 residents distributed across six rural counties in Missouri. The facility offers inpatient care including surgical and obstetric services, and outpatient surgical and cardiac rehabilitation. The hospital employs 225 staff and providers (compared to urban medical centers which employ on average 1500+) and has been operating its physical plant at its current location for 9½ years. Nearly 22% of patients seen at the hospital are 65 years or older. The number of patients seen at the hospital between the ages of 0 (birth) and 17 years is 25%. Approximately 15% of the area's populations served have income below the poverty level.

Rural hospitals such as Midwest continually encounter a challenging financial landscape with decreasing demand for inpatient services, revenue pressure, complex patient populations and the cost of attracting and retaining medical providers. Between 2010 and 2021, 138 rural hospitals in the U.S. closed their doors, further reducing the availability of services to their communities (NC Rural Health Research, 2022).

In order to help Midwest Hospital retain some of its service lines and develop additional programs, the organization received a grant in 2018 through the Health Resources & Services Administration to support its *Aging in Place* program. Due to the large population of older adults who would prefer to age in place (remain in their homes and communities for as long as possible) in its service area, the hospital has been seeking to provide enhanced medical care to senior citizens. Grant funds were directed toward increasing home health services, improving technology for telehealth, and increasing preventive care.

Following a reduction in COVID-19 hospital admissions, Midwest Hospital has again begun analyzing the availability of services and needs for its regional populations. Midwest's administration is contemplating whether to apply for another grant, this time to support children's

vision services. Midwest currently employs two optometrists. Family practitioners at the hospital and associated clinics also participate in the vision screening of children; and all providers have noticed an uptick in vision problems—especially myopia (nearsightedness)—in child visits beginning late 2020. (With myopia, objects that are close-up look clear but distant objects look blurry.) Both Midwest administrators and medical providers are interested in determining if this increase in vision problems is an anomaly or will be a continuing issue for the future. If a persistent problem, the hospital would need to consider hiring additional optometric staff and expanding screening services and facilities.

Expansion of vision services would additionally benefit the hospital's Aging in Place program. Age-related macular degeneration has also been noted as rising in incidence in older patients. This vision loss makes it difficult for individuals to drive, watch television, read, and complete household tasks such as cooking or cleaning (NEI, 2022). Midwest could also investigate the establishment of an ophthalmic telehealth program for the treatment of eye disease, which would benefit both adults and children.

The first order of business in preparing the grant application and planning for expanding vision services is to determine need. The hospital has logged a record number of myopia cases in children (aged 0-17 years) from the last quarter in 2020 through the first quarter of 2022. Working with an urban medical center librarian, Midwest Hospital providers uncovered a disturbing pattern in children's vision beginning in 2020. Midwest has asked their sole health data analyst to answer questions about this trend using data—from regional and national sources — while they research the literature.

3. WHAT IS SEEN IN THE LITERATURE

During their research, the providers found that myopia and associated vision loss has been recognized as a significant issue for some time. The World Health Organization estimated in 2015 that half of the world's population may be myopic by 2050 (Holden et al., 2015). The literature they reviewed demonstrated that myopia was rising even more rapidly than expected throughout the world.

A recent Chinese study of over 195,000 school-aged children noted a significant myopic shift in the year 2020 compared to previously recorded years (2015-2019). Specifically, the study found that the prevalence of myopia in the 2020 screenings was higher than the highest prevalence of myopia within 2015-2019 for children aged 6 (21.5% vs 5.7%), 7 (26.2% vs 16.2%), and 8 (37.2% vs 27.7%) years (Wang et al., 2021). Researchers theorized that the at-home confinement of children during the pandemic, combined with insufficient outside activities and increased screen time, were associated with the myopic changes. The medical center librarian located dozens of similar studies over the period 2020 through 2022, all suggesting that isolation due to COVID-19 was responsible for much of the observed increases in myopia cases.

Although myopia is generally regarded as a benign condition that is treatable with corrective lenses and other therapies, untreated cases are associated with an increased risk of blinding pathologies such as cataracts and glaucoma (Zhang et al., 2022). Early detection through screening is the key to prevention of these severe progressions, and also aids children with early correction so they have improved sight quality.

Midwest Hospital providers were both relieved and concerned by information they discovered. They were relieved in the knowledge that what they were experiencing in their small rural hospital population was not an isolated situation; yet concerned about the broader implications for their young patients. After receiving the librarian's report, Midwest administrators turned to their health informaticist to evaluate results of the data analysis.

4. THE BIG PICTURE IN NUMBERS

The data file references an actual children's vision and eye health surveillance study organized by the Maternal and Child Health Bureau of the U.S. Department of Health and Human Services. The data set contains demographic information that identifies the gender, race/ethnicity, and location of children being screened for vision impairment or issues. The records cover simulated surveillance results for the years 2016 through 2017 (pre-COVID-19 pandemic) with similar data for the years 2020 through 2021 (post the start of the COVID-19 pandemic). All the data represents individuals in the 0-17 age group in all 50 states. The data file for this exercise contains simulated records and should only be used for this exercise. The data file will be maintained at: https://drive.google.com/file/d/1Hx31CRh9oCJFMQcIx-fBt-8gXpNrBcOy/view?usp=sharing

Using the exercise data file and the data dictionary provided in Appendix A, the analyst is asked to see if they can answer the questions provided in Appendix B for the hospital's grant application through a comparison of the 2016-17 (1,968 records) and 2020-21 (2,400 records) sets of data).

After the analyst has examined the data, Midwest administration has asked that the results be reported in a table which may be sent to the grant-writer for further analysis and preparation of the grant application. A summarization of the analysis should be included with the table and charts effectively answering the questions posed in Appendix B. No conclusions or measures of significance have been requested on the part of the analyst at this time.

5. BIBLIOGRAPHY

Botelho, B. & Bigelow, S. J. (2022). Big Data. *Tech Accelerator.* Retrieved May 5, 2022, from https://www.techtarget.com/searchdatamanagement/definition/big-data

HHS ASPR TRACIE. (2020). Rural Health and COVID-19. Retrieved May 4, 2022, from https://files.asprtracie.hhs.gov/documents/aspr-tracie-rural-health-and-covid-19.pdf

Holden, B., Mariotti, S., Kocur, I., Resnikoff, S., He, M., Naidoo, K., et al. (2015). The impact of myopia and high myopia: Report of the joint World Health Organization- Brien Holden Vision Institute Global Scientific Meeting on Myopia University of New South Wales, Sydney, Australia, 16–18 March 2015. Geneva: World Health Organization.

Jafar, M., Anderson, R., & A. Abdullat. (2008). Data mining methods course for computer information systems students. *Information Systems Education Journal.* *6*(48), http://isedj.org/6/48/. ISSN: 1545-679X

Kent, J. (2021). Collecting big data to eliminate rural health disparities. *Health IT Analytics.* Retrieved April 20, 2022, from https://healthitanalytics.com/news/collecting-big-data-to-eliminate-rural-health-disparities

National Eye Institute. (2021). Age-related macular degeneration. *Learn about Eye Health*. Retrieved May 4, 2022, from https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/age-related-macular-degeneration

NC Rural Health Research Program. (2022). Rural hospital closures maps, 2005- Present. – The Cecil G. Sheps Center for Health Services Research. Retrieved May 5, 2022 from https://www.shepscenter.unc.edu/programs-projects/rural-health/rural-hospital-closures/

Pastorino, R., De Vito, C., Migliara, G., Glocker, K., Binenbaum, I., Ricciardi, W., & Boccia, S. (2019). Benefits and challenges of Big Data in healthcare: An overview of the European initiatives. *European Journal of Public Health*, *29*(Supplement_3), 23–27. https://doi.org/10.1093/eurpub/ckz168

Wang, J., Li, Y., Musch, D. C., Wei, N., Qi, X., Ding, G., Li, X., Li, J., Song, L., Zhang, Y., Ning, Y., Zeng, X., Hua, N., Li, S., & Qian, X. (2021). Progression of myopia in school-Aged children after COVID-19 home confinement. *JAMA Ophthalmology*, *139*(3), 293–300. https://doi.org/10.1001/jamaophthalmol.2020.6239

Yoo, I., Alafaireet, P., Marinov, M., Pena-Hernandez, K., Gopidi, R., Chanage, J. and L. Hua. (2012). Data mining in healthcare and biomedicine: A survey of literature. *Journal of Medical Systems*. 26. 2431-2448.

Zhang, C., Li, L., Jan, C., Li, X., & Qu, J. (2022). Association of school education with eyesight among children and adolescents. *JAMA Network Open*, *5*(4), e229545. https://doi.org/10.1001/jamanetworkopen.2022.9545

Editor's Note:

This paper was selected for inclusion in the journal as an EDSIGCON 2022 Distinguished Case. The acceptance rate is typically 7% for this category of teaching cases based on blind reviews from six or more peers including three or more former best case authors who did not submit a paper in 2022.

APPENDIX A
Data Dictionary for Vision and Eye Health Surveillance Data File

Fields / (Field Values)	Description for Fields and Some Field Values	Data Type
YearStart	Starting Year for year range	number
YearEnd	Ending Year for year range, same as starting year if single year used in evaluation	number
LocationAbbr	State Abbreviation	plain text
LocationDesc	State Name	plain text
DataSource	Abbreviation of Data Source	plain text
Topic	Topic Description (Values defined below)	plain text
(BLND)	Blind or severe difficulty seeing	
(DGLAS)	Difficulty seeing with glasses	
(MYOPIA)	Near-sightedness (MYOPIA)	
(HPYEROPIA)	Far-sightedness (HYPEROPIA)	
Category	Category Description	plain text
Question	Question Description	plain text
Age	Stratification value for age group e.g. 0-17yrs	plain text
Gender	Stratification value for gender e.g. Male, Female	plain text
RaceEthnicity	Stratification value for race e.g. White, non-Hispanic	plain text
RiskFactor	Stratification value for major risk factor e.g. diabetes	plain text
Insurance	Type of Insurance (Values defined below)	plain text
(Ins_D)	Medicare + Medicaid Dual Eligible	
(Ins_E)	Medicaid	
(Ins_S)	Medicare Fee for Service	
(Ins_C)	Medicare Managed	
(Ins_Y)	Military	
(Ins_G)	Other Government	
(Ins_P)	Private	
(Ins_U)	No payment listed	
(Ins_All)	All payers	
Designation	Metropolitan Statistics Area Designation (Values defined below)	plain text
(URBN)	Urban: within a principal city of a MSA	
(SUBN)	Suburban: w/in a MSA but not w/in a principal city of the MSA	
(RURL)	Rural: outside of an MSA	
Diagnosis ID	Diagnosis of vision issue (Values defined below)	plain text
(BLND)	Blind or severe difficulty seeing - (Visual function)	
(DGLAS)	Difficulty seeing with glasses - (Visual function)	
(MYOPIA)	Near-sightedness (MYOPIA) - (Visual function)	
(HPYEROPIA)	Far-sightedness (HYPEROPIA) - (Visual function)	

(CMYOPIA)	Refractive correction for Myopia- (Corrective procedure)	
(CHYPEROPIA)	Refractive correction for Hyperopia - (Corrective procedure)	
(COTHER)	Refractive correction for Other - (Corrective procedure)	
(CATARACT)	Cataract surgery - (Corrective procedure)	
GeoLocation	Geographic location	plain text

APPENDIX B
Data Analysis Questions

1. What was the overall change in the following *Visual Function* categories in participants aged 0-17: Blind or Difficulty Seeing (BLND); Difficulty seeing with glasses (DGLAS); Near-sightedness (MYOPIA); and /or Far-sightedness (HYPEROPIA)?
2. What was the overall change in the following *Visual Function* categories in participants aged 0-17 by gender: Blind or Difficulty Seeing (BLND); Difficulty seeing with glasses (DGLAS); Near-sightedness (MYOPIA); and /or Far-sightedness (HYPEROPIA)?
3. What was the overall change in the following *Visual Function* categories in participants aged 0-17 by race/ethnicity: Blind or Difficulty Seeing (BLND); Difficulty seeing with glasses (DGLAS); Near-sightedness (MYOPIA); and /or Far-sightedness (HYPEROPIA)?
4. What was the overall change in the following *Service Utilization* categories in participants aged 0-17: Refractive correction for Myopia (CMYOPIA); Refractive correction for Hyperopia (CHYPEROPIA); Refractive correction for Other (COTHER); and/or Cataract surgery (CATARACT)?
5. What was the overall change in the following *Service Utilization* categories in participants aged 0-17 by gender: Refractive correction for Myopia (CMYOPIA); Refractive correction for Hyperopia (CHYPEROPIA); Refractive correction for Other (COTHER); and/or Cataract surgery (CATARACT)?
6. What was the overall change in the following *Service Utilization* categories in participants aged 0-17 by race/ethnicity: Refractive correction for Myopia (CMYOPIA); Refractive correction for Hyperopia (CHYPEROPIA); Refractive correction for Other (COTHER); and/or Cataract surgery (CATARACT)?
7. What was the overall change in the following *Visual Function* categories in participants aged 0-17: Blind or Difficulty Seeing (BLND); Difficulty seeing with glasses (DGLAS); Near-sightedness (MYOPIA); and /or Far-sightedness (HYPEROPIA) in these states: Missouri? Kansas? Illinois? Arkansas?
8. What was the overall change in the following *Service Utilization* categories in participants aged 0-17: Refractive correction for Myopia (CMYOPIA); Refractive correction for Hyperopia (CHYPEROPIA); Refractive correction for Other (COTHER); and/or Cataract surgery (CATARACT) in these states: Missouri? Kansas? Illinois? Arkansas?
9. Of participants aged 0-17, what was the overall change in the combined *Visual Function* category Near-sightedness (MYOPIA) and the *Service Utilization* category Refractive correction for Myopia (CMYOPIA)?
10. How many participants aged 0-17 in the combined *Visual Function* category Near-sightedness (MYOPIA) and in the *Service Utilization* category Refractive correction for Myopia (CMYOPIA) lived in rural communities?
11. Of participants aged 0-17, what was the overall change in the combined *Visual Function* category Near-sightedness (MYOPIA) and the *Service Utilization* category Refractive correction for Myopia (CMYOPIA) in these states: Missouri? Kansas? Illinois? Arkansas?
12. How many participants aged 0-17 in the survey were insured by Medicaid? How many were uninsured? How many had private insurance?
13. How many rural participants aged 0-17 in the survey were insured by Medicaid? How many were uninsured? How many had private insurance?
14.

Perceptions of IT students' utilization of embedded online tutors

S.C. Spangler
scott.spangler@mga.edu

Kamy K Shah
kamy.shah@mga.edu

Wayne E. Lockwood
wayne.lockwood@mga.edu

Department of Information Technology
Middle Georgia State University
Macon, Georgia 31206 USA

Abstract

The pilot study reflects perceptions from higher education students in an experimental new online teaching program at a mid-sized Southeastern United States University. The research focused on the effectiveness of an embedded tutoring pilot program in online and hybrid learning management systems (LMS). The research was focused on information technology students' (n=46) perceptions of comfort, confidence, and utilization of the pilot program to understand its value in student retention. The research notes student comfort, confidence, and utilization of the program. Additionally, it supports the notion that the program can construct student retention aspects by reducing anxieties and stress from distance learning spatial inconsistencies. The findings suggest similar results or parallel considerations of students' perceptions in the literature on embedded librarian programs. However, the results fail to mirror students receiving more significant levels of self-efficacy.

Keywords: Embedded tutor, course-integrated supplementation, online IT instruction.

1. INTRODUCTION

Today, the world is still at war with the Corona Virus. The force of the virus's nature has forever changed all aspects of everyday life, including our education systems. Faced with unknown limitations and anxieties, Universities and k-12 programs across the globe were forced to change their pedagogical methodologies to influence business continuity.

The classroom mechanics are simply sustained through innovative information systems channel changes. Silva (2015) explains the speed at which society's forces create technology acceptance and institutes immediate and direct modification in an information system is conditional by human organizations unless under duress of high priority in all spheres of society.

Early adopters and the conditions under which humans or organizations adopt information systems are still a high priority of scientific research and vastly undetermined (Venkatesh & Davis, 2000). Initially, Rogers' (2010) stated that innovations "require a lengthy period, often of many years, from the time they become available to the time they are widely adopted" (p. 1).

Nevertheless, past research claims universities sustained an extraordinarily high rate of acceptance and speed in the acceptance of rectifications to the way teaching was conducted. And to enhance their information systems through changes and adoptions, faculty were forced to recognize the immediacy of students' needs and have empathy. To increase comfort and confidence and reduce anxiety, faculty continuously attempted to introduce new online innovations in the classroom to students that are currently recognized as retention success factors, like the embedded librarian programs. The embedded librarian programs entrench library staff members into online courses as contributors to curriculum instruction and support, offering immediate student assistance (Spangler et al., 2020). Shadowing the success levels of the embedded librarian program, new information systems models have sprung into the Learning Management Systems (LMS) modeling to further navigate difficulties in students' transitions to online education. This paper will seek to understand students' perceptions of one new information system channel in the LMS design–embedded tutors.

2. LITERATURE

Evolving Offerings in LMS

The belief that there is a more efficient process to learning motivates many to seek new innovative ways and tools. Scholars continue to study early adopters on how, why, and under what conditions innovation and new technology can and will be used (Venkatesh & Davis, 2000). Innovative technologies provide many amenities that traditional courses may not offer.

Scholars have argued about technology use in the classroom as far back as the invention of the teaching machine by Sidney Pressey (Petrina, 2004). The "Automatic Teacher" was designed to automate testing by letting students assess themselves and was considered the first Learning Management System (LMS) developed. Although the machine was a success, the concept was never socially accepted because of the lack of commitment by other scholars and was never officially used as a proper tool (Petrina, 2004).

Early adoption of an innovation or new technology was often decided by demographics and luxury investment rather than essential needs. Reardon et al. (2019) stated that many variables could contribute to a successful education. Technology can be difficult to acquire if school budgets are low and if they reside in low socioeconomics, they are not always treated equally (Allen, 2019). Economics, demographic, segregation, and school opportunities are crucial in attaining innovative technology. White (2019) stated that students with special needs, color, and low income suffer inequities and are less likely to acquire technology without pressure or additional support.

Many researchers argue that technology is not always the answer and can be deconstructive rather than constructive. Turkle (2011) also stated that technology addiction and misuse are often asserted when implementing technology. Turkle (2011) asserted that the phenomenon of smartphone obsession is argued as a nuisance in society and creates a social disconnect and promotes a breakdown in personal communication. Sukenick (2012) agrees that technology minimizes interaction and suppresses interactions. Spangler (2015) argued that society has become desensitized to technology's tribulations and has become riddled with anxiety, nervousness, and fear of disconnect to the point of disillusionment.

Teachers are learning to embrace and benefit from innovative technology to support and enhance pedagogy (Martin et al., 2019). Innovative technology has decreased the literacy gap, increased student retention, and opened doors for underprivileged or disenfranchised (poor or geographically challenged) students. It has additionally allowed virtual courses to exist that would otherwise not be possible, increasing the accessibility of higher quality education to students. Access and limitations of schools caused by the presence of Covid-19 have demonstrated that innovative technology is an asset that all schools need. Nowicki (2020) stated that the Covid-19 pandemic demonstrated that many schools were not fully prepared for a disaster recovery plan for long-term instruction. Other concerns are the disconnect between students from both online instruction and administrators. There are various challenges when implementing and utilizing innovative technology in a classroom, especially in rural and low socioeconomic areas (Carr-Chellman et al., 2020). The acceptance of LMS provides a valid option.

Technology advancements in learning environments differ and provide multiple arguments for the best results. Finding the appropriate balance for the student can be challenging. An example is embedded library instruction. Some schools are responding to the detachment and physical separation often felt by students and faculty from asynchronous classes

caused by the attributes of online and distance education innovations (Spangler, 2020). Spangler (2020) stated that innovative technology using real-time or near real-time metric instruments facilitated by faculty members via LMS provides support for asynchronous online learning. Spangler argues that these innovative embedded apps reduce distance learning anxieties. According to Garcia-Castelan et al. (2021) and Spencer and Temple (2021), traditional teaching, also referred to as brick and mortar or face-to-face (F2F), provides many benefits over online teaching. Garcia-Castelan et al. (2021) stated that many students prefer the F2F learning process because it promotes teacher and group activities. Spencer and Temple (2021) found that students' performance and attention to detail are more focused when learning F2F.

Educational learning platforms are continuing to evolve. LMSs are a critical platform for online approaches such as blended, web-based, and distance learning and provide an extension to the traditional F2F class experience. Many students do not have the opportunity to attend traditional F2F classes because of lifestyle obstacles. Innovative technology provides many amenities that traditional courses may not offer. To evade life obstacles, students seek a more favorable learning method that can offer the flexibility of time restrictions, easy access, and user-friendly environments. According to Spangler (2019) and Spencer & Temple (2021), LMS technology advancements provide a user-friendly medium that centralizes the academic workspace via the internet, allowing students to utilize personal technological devices while supporting academic learning.

Learning management systems were designed to support distance learning by connecting multiple students in different geological areas forming a virtual classroom (Tumbleson, 2016; Spangler, 2019; & Spangler et al., 2020). According to Tumbleson (2016) and Spangler (2019), the programs increase awareness and exposure to embedded librarian resources and lead to improved library resource utilization and support for students in LMS-based courses. Embedding librarians directly into LMS provides essential support for educators and creates a virtual library and support liaisons for the student. Tumbleson (2016) stated that implementing the embedded librarian and support may vary and is primarily controlled by faculty. Tumbleson (2016) explained that some university libraries adopt a macro process using relevant links and subject templates to focus on the relevant subject using a LibGuide, research template or post a link to the university library website URL with contact information. However, with the collaboration of faculty members, Tumbleson clarifies that students can be provided a collaborative approach through shared resources where the Librarian seamlessly interacts with the student on course-related research assignments. According to Tumbleson, most LMSs provide customizable features to create embedded librarian pages and widgets that offer easy quick-link contact resource information.

Educators could make use of this easy-to-build LMS addition to a course. Educators understand the findings could increase student retention in their courses and add a level of comfort in the teaching efforts. Spangler's (2019) research noted students embraced the program. The researcher found the online connection to provide a level of self-efficacy-building attributes and overall academic confidence in abilities to succeed in a course. Additionally, educators should note the embedded programs in the online learning management systems' courses offered a perception of confidence building, reduction in anxieties, and self-efficacy constructions for students. The program and past literature support the positive effects online students receive from having fingertip connectivity to distance learning support additives in learning management systems.

The Role of Embedded Librarians During a Global Pandemic

The limited research in the area focuses on the empowerment of embedded programs and how they offer student support, anxiety reduction, confidence building, and self-efficacy perceptions. The immediacy for student support in distance learning was realized because of the disenfranchised students during the global pandemic. All students were immersed in online learning modalities and left stranded for traditional services. The literature focus originates in the scholars Edwards et al. (2010), that discovered the power of embedding a librarian in a hybrid style course. The research presented a new level of confidence, comfort, and anxiety reduction. This research spurred many notes to educators about the importance of utilizing services in LMS-supported courses when online services originally became conceptualized. Although the pilot program had attributes of face-to-face connectivity, the LMS supported the information transfer to the students only as a mechanism to foster disenfranchised distance students' access to resources. This finding was similarly concluded in later research directed entirely on the online students' perceptions of

embedded resources such as the librarians (Spangler, 2019; Spangler et al., 2020; and Spangler, 2020).

Edward et al. (2010) described the term embedded as being borrowed from the "practice of embedding journalists in combat zones during military conflicts and refers to complete integration" (p. 273). The online embedded support "supplement" creates ease for students to navigate finding resources in a distance learning situation. The goal of Spangler's (2019) research sought first to understand the students' perceptions of the embedded pilot program. And secondly, the research seeks to understand if the students' perceptions are congruent with literature on embedded librarians' effects in an online course: confidence, comfort, anxiety reduction, and self-efficacy building in students. The findings concluded that the embedded programs offered to students in face-to-face and hybrid courses were congruent to Edward's et al. findings and levels of confidence building.

Interestingly noted, the students' perceptions from both pilot programs ran congruent. Students in Spangler's (2019) research remarked the fingertip assistance of the embedded aspects created self-efficacy and a willingness for students to stay in a course rather than withdraw. The embedded programs build confidence in research abilities, assignments, and anxiety reduction from feeling detached from a traditional university setting with face-to-face services. Research from Spangler et al. (2020) highlighted the positive perceptions of having an embedded librarian in an online course created considerable amounts of student confidence. Additionally, Spangler's (2020) research on graduate students found that embedded online programs and assistance increase student comfort and confidence in their abilities to research and complete complex assignments.

Embedded librarians are a modern innovation to facilitate a greater experiential learning environment for students taking online university courses–either partially (hybrid) or completely (Alsuqaih, 2020). Spangler et al. (2020) concluded that the "students' perceptions of embedded librarians in online and hybrid courses" are positive. The pilot study on embedded programs establishes that students' perceptions of online experiences offer a higher learning experience with "comfort, interpersonal, academic confidence, confidence in abilities, and confidence in researching and citing sources for assignments" (p. 173). Similarly, Matteson's (2020) research on students in virtual classes concluded a strong need for embedded online university services. The scholar concluded that the university's need for embedded services would grow significantly post-pandemic, and higher education pandemic models will continue to change education globally (Matteson, 2020).

Furthermore, the scholar suggested that "librarians are quickly moving their instruction online to manage the restrictions of social distancing because some or all of their students are learning remotely" and to diminish the bottleneck effect that instructors or tutors might experience (Matteson, 2020, p. 24). Congruent in findings, Steele's (2021) research further forecasts the necessity of the new online services provided by embedded librarians–because of the COVID-19 pandemic changes in educational pedagogy. The University of Southern Mississippi researchers concluded that the expanding growth of virtual students–post-pandemic–is why universities need to further programs and increase dynamic support methods for students.

Embedded Librarianship Assisting Skilled Concentrations

The Guillot et al. (2010) study analyzes the relationship between nursing faculty and embedded librarians. The researchers stated that students must maintain a relationship with the subject, faculty member, and their embedded instructor for the most direct assistance or relevance. Spangler et al. (2020) offered this finding and suggested that embedded librarians can offer students confidence, comfort, and self-efficacy from the solid relationships and trust built from the programs. Similarly, Franzen and Sharkey (2021) examined the impact of the services of embedded librarians on undergraduate nursing students and their ability to configure information skills. Franzen and Sharkey (2021) study stated that embedded programs typically or "rarely have a long-term impact on students' research behavior or skill sets" with the standard in-person library sessions (Franzen & Sharkey, 2021, p. 311). Additionally, recent research concludes that embedded librarians are unnecessary for students (Wu et al., 2021), especially those studying nursing. Wu et al. hypothesized in the research that specialized fields of concentration whereby embedded librarianship do not have a focus on the studies nor applications (like the nursing or general health care practitioner fields) yield difficulties for librarians to foster student success and measure of value.

Other research contrasts Wu et al. (2021) considerations suggesting a greater need for embedded librarians with specific concentration

knowledge in developing student relationships too (Menard & Misquith, 2021). Additionally, Chan's (2021) study findings counter the argument that embedded librarianships can offer nursing students numerous benefits. Specifically, the researchers observed that students interacting with field-focused nursing-specific embedded librarians could assist nursing students in improving (Franzen & Sharkey, 2021, p. 311).

Students' Perceptions of Embedded Tutor Services in Online Higher Education Courses

Shumaker (2014) stated that traditional in-person university professionals "need to be fully "read into" the nature of the work being performed" and "need a full understanding of the nature of the task and the goals of the effort" to accomplish the objectives of the services to gain student satisfaction and confidences (p. 5). The scholar suggests that virtually embedded "relationships" require trust and partnership between the faculty and liaison to incorporate support into the curriculum properly. Embedded tutors are a modern innovation to facilitate a more experiential learning environment for those who have university online courses (Alsuqaih, 2020). Spangler et al. (2020) research "demonstrated courses with embedded programs offer students a 'higher rate of comfort, interpersonal, academic confidence, confidence in abilities, and confidence in researching and citing sources for assignments'" (p. 173).

Similarly, Mendoza and Kerl (2021) research uncovered students perceive online embedded tutoring as being "meaningful" to users and further the "academic and social benefits" of online learning (p. 69). Other scholars suggested that the pandemic forced immediate changes in online services to students, creating new avenues for experimental collaborative adventures in pedagogy with new grounds for software supplementation and training (O'Brien, 2020; Sonn et al., 2021). Mendoza and Kerl's case study determined which tactics for integrating tutors and other resources throughout online courses have been employed. The authors even affirm that embedded tutors can be multi-functional and express that "governments should allocate funding to employ more qualified teaching tutors to assist lecturers with the teaching and grading workload" (Sonn et al., 2021, p. 12). Matterson (2020) furthered this regard by noting embedding tutors in online courses allows academic service managers to "manage" social distancing or "learning remotely" to diminish the bottleneck effect instructors, librarians, or specialists might experience (p. 24).

To build retention in classrooms and successful strategies, faculty must incorporate embedded tutors' viewpoints into their curriculum (Shumaker, 2014). The research suggests that when faculty build curriculum and assessment material with embedded tutors, the process creates trust, establishes tutors' roles in learning, and strong collaboration for student success. Mendoza and Kerl (2021) suggested that embedded tutors need to have an essence of self-efficacy or buy-in when developing and working with faculty in an embedded tutor program on the macro level. The researchers' study questioned how embedded tutors could foster student self-efficacy through online course services. The researchers concluded that "the use of embedded tutors can be a beneficial approach for learning" to support students (p. 69). But the researchers noted that "studies are needed to explore further the academic and social benefits of embedded online tutors and training" (p. 69). Their findings determined that the services are ultimately meaningful to the users. Still, future studies are needed to determine the perceptions of students and the enhancement embedded tutors in an online course can yield for determining learning outcome value.

3. METHODOLOGY

An IRB-approved survey instrument was administered to a small pilot population (n=46) of undergraduate and graduate students attending information technology courses at Middle Georgia State University. With permission from past research by Spangler (2020), the pilot instrument was designed to understand students' perceptions of the new embedded tutoring program in information technology courses. The study was conducted to understand the students' perceptions of a new pilot program. The program's focus and intentions were to recognize the support an online embedded tutoring program could have on student success and course retention. Additionally, the study focus pondered if the new pilot program addition to an online course would have similar findings of comfort, confidence building, and anxiety reduction as its parallel program uncovered by having an embedded librarian.

The survey instrument was administered through a Google Form after the conclusion of face-to-face and online courses to reduce population bias and increase validation. The survey was not tracked and required consent before starting the research. Any participant under 18 years old was not permitted to participate. All participants completed the instrument entirely. The after-

course assessment allowed participants to have reduced measures of worry about grade conflicts from open comments and the overall value assessment of the embedded tutoring program. Additionally, open comments were allowed to capture additional notes of participants' thoughts. To further create validity, the researcher allowed an external reviewer to verify the data's findings and the researchers' conclusions. Overall, the research focused on aspects of synchronized tutoring learning supplements to an online information technology course.

The pilot program was an addition to two online information technology project management courses. Both courses had multiple supplemented online attributes, including an embedded librarian and an embedded tutor. The instrument was administered to online distance learning management system sections and one hybrid model. Students were alerted to the pilot program during the instructor's introductory course video overview. The embedded tutor's contact information and scheduling link were also inserted into the LMS' home page as a sidebar widget application. The easy-to-find widget's construction allowed students to click on the embedded tutor's image and resource link to immediately contact the tutor and Student Success Center's helpline for scheduling. Additionally, the embedded tutor constructed a "how to schedule" appointment and prepared for the tutoring session video. The video link was located inside the LMS course homepage widget, allowing students to watch and understand the program's services easily. Students were directly linked upon completing the instructions to an online scheduling tool and instructional video for learning about online interfacing for tutoring.

RQ1: Do students perceive embedded tutoring programs as offering educational benefits?

4. FINDINGS

The 23-question instrument received 46 complete responses out of the 53 requested responders. The instrument received no incomplete responses. Hence, no responses were eliminated from the analysis to forecast a complete representation in the pilot study. All participants were asked to sign a voluntary consent form before completing the instrument. The population received a nearly equal value in gender (male n=59% and female n=41%). Most participants were information technology students (n=61%); however, some responders were from other majors as the courses were available to any significant (business n=13%, School of Arts and Letters n=5%, School of Aviation n=2%, School of Health and Natural Sciences n=13%, School of Graduate Studies 5%, and School of Education and Behavioral Studies 1%).

The instrument first assessed participant utilization of university success services to understand the responders' acceptance level, knowledge, and comfortability with university services. Participants noted a relatively neutral level of acceptance for utilizing the university's Student Success Center (24% extremely unlikely, 17% unlikely, 18% neutral, 19% Extremely likely, 22%). Responders feel "neutral" by stating they had abilities to use the Success Center's tools (13% extremely unlikely, 24% unlikely, 31% neutral, likely, 17% Extremely likely, 15%). And interestingly noted, the participants noted similar neutral considerations about asking for research help at a Student Success Center (24% extremely unlikely, 11% unlikely, 19% neutral, likely, 22% Extremely likely, 24%). Although, the population considered it a greater comfort (n= 50%) to use the Student Success Center for virtual tutoring and mentoring (24% extremely unlikely, 13% unlikely, 13% neutral, likely 25% Extremely likely, 25%). But interestingly noted, the traditional tutoring methods (not online) from the center remained neutral (22% extremely unlikely, 16% unlikely, 18% neutral, likely, 20% Extremely likely, 24%). Overall, this noted a polarized perception of using the Success Center.

Open Educational Resources and Tutors Tools

Interestingly, participants were again neutral about accepting and using free Student Success Center's instruments (37% unlikely, 13 % neutral, 50% likely) designed around open educational resources and asking for help in finding these instruments (44% unlikely, 13 % neutral, 43% likely). Additionally, the participants had a neutral desire to understand how to use the tutor's open educational resource (OER) designed tools for learning APA Style correctly (19% extremely unlikely, 24% unlikely, 13% neutral, likely, 13% Extremely likely, and 31%). Although the majority did find value in using the tutor's OER tools to avoid plagiarism (24% unlikely, 18% neutral, 58% likely).

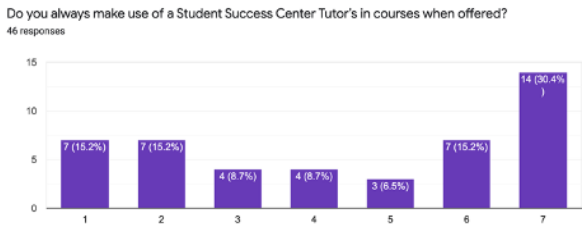


Figure 1: *Participants' acceptance level of embedded tutors*

When directly questioned about perception levels of the new embedded tutoring program in online and hybrid courses (Figure 1), participants observed a (45%) level of acceptance. And a similar level of acceptance (45%) to use the embedded tutor for research help and again for virtual assistance on any subject matter (45%). Interestingly, participants found a negative level of motivation (60%) or self-efficacy gains from having an embedded tutor in the courses. Because of these main findings, the scholars cannot state a clear and robust benefit and recognize the limited success of RQ1: Do students perceive embedded tutoring programs as offering educational benefits? However, the scholars' findings note that half of the population did feel that without the embedded tutor in the course, more participants would be likely to withdraw from a course (50%) although they may not use the services.

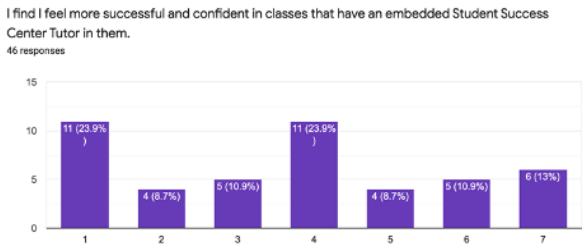


Figure 2: Participants' confidence with embedded tutors in classes

Nevertheless, the self-efficacy levels of the participants reflected a slight confidence level boost (53%) from observing the tutoring of the course learning management (LMS) shells. And they again observed a little confidence (37% unlikely, 13 neutral, 50% likely) for receiving a higher level of achievement on assignments from the services engaged. Interestingly noted, participants regarded their confidence levels (Figure 2) as not increasing dramatically from the presence of an embedded tutor in the hybrid or online course (LMS) shells (24% extremely unlikely, 9% mostly unlikely, 11% unlikely, 23% neutral, likely 9%, most likely 11%, extremely likely 13%). Straightaway, when asked if an embedded tutor's photograph listed in the (LMS) online would offer comfort or confidence, participants found no value (75% of the population states the image is unlikely to have an impact).

5. DISCUSSION

The researchers first acknowledge the low population in the pilot study program cannot be generalized to a larger population. Hence, the findings here in the discussion are limited in scope and overall determination. Secondly, the researchers note the population perceives the program as having limited merits of educational benefits suggesting RQ1 is not fully supported. Nevertheless, the research does suggest a universal academic need to create avenues and methods for student success in online and hybrid courses in the future, which was noted in prior research (Alsuqaih, 2020; Spangler et al., 2020; O'Brien, 2020; Sonn et al., 2021). The pilot program started at the height of the pandemic and introduced a new line of technology integration and innovations in online higher education courses.

Slightly over half the population found value in the pilot embedded tutoring program. Nearly half of the participants found value in the program, which suggests a need for embedded tutoring options. Interestingly, the respondents showed similar findings to past research on embedded librarian programs (O'Brien, 2020; Spangler et al., 2020; Sonn et al., 2021). In this pilot study, participants seemed polarized. Half of the respondents did agree with the past embedded librarian research that the virtual aspect increases confidence (50% likely) and comfort levels (50% likely). But most importantly noted, respondents stated the tutoring program could create motivation in online students (60% likely). This regard was mirrored in Mendoza and Kerl's (2021) research, suggesting students found the programs "meaningful" (p. 69). Mendoza and Kerl's research suggested that embedded tutors need an essence of buy-in when developing and working with faculty in an embedded tutor program on the macro level. This buy-in effect creates a greater essence in the classes and allows students to regard the program as valuable to their academic journeys.

However, in other research, online images of embedded librarians offered greater comfort, self-efficacy, and regard for having success in an online course (Spangler et al., 2020). Interestingly noted in this study,

participants' perceptions demonstrated no real value with the virtual photographs of an embedded tutor in the courses (75% of the population find the program has no value). This note may be directly related to the fact most of the embedded tools being used in the LMS were driven by the embedded librarian's program and not the embedded tutor, who focused on teaching students how to utilize the instruments to the height of effectiveness. Additionally, self-efficacy levels were demonstrated higher in other studies (Spangler et al., 2020; Mendoza and Kerl, 2021; Sonnet al., 2021), suggesting a need for deeper understanding and further understanding research into student's self-efficacy levels in online courses offering embedded programs.

The researchers suggest creating a more extensive and diverse population to further this study. Furthermore, the researchers suggest cross-examining the instrument against populations outside of the information technology field of study, whereby the students are vested in technology resources, instruments, and tools widely and prolifically in their natural course study habits.

6. CONCLUSION

The research was focused on information technology students' (n=46) perceptions of comfort, confidence, and utilization of the pilot embedded tutoring program. Secondly, the paper sought to understand if the pilot program could create student retention. The researchers' observations can't be fully supported or generalized to state that the embedded tutoring program is perceived to have a propensity for online student retention. However, the self-efficacy levels of the participants did reflect a slight confidence level boost (53%) from observing the tutoring option inside the course learning management (LMS) shells. Additionally, the research noted positive student course comfort, assignment confidence, and a slight emphasis on self-efficacy from program utilization. Unfortunately, the findings on embedded tutors are not as merit worthy as other embedded programs like the use of embedded librarians (Spangler, 2020; Spangler et al., 2020).

Lastly, the researchers note the population perceives the program as having merits of educational benefits suggesting RQ1 is supported, but from only slightly over half of the population. And only half of the respondents found the embedded tutoring program merit worthy and successful in aiding their academic journey. Nevertheless, despite the polarized findings, the research can conclude that embedded programs are needed in LMS course shells for online students' benefits. According to the findings, the researchers can conclude that students perceive embedded tutors as having some value in their overall education experience. The research supports the notion that the embedded tutoring program can construct student retention aspects by reducing anxieties and stress from distance learning spatial inconsistencies for at least half of its population. Nevertheless, the research does suggest a universal academic need to create avenues and methods for student success in online and hybrid courses in the future, which was noted in prior literature (Alsuqaih, 2020; Spangler et al., 2020; O'Brien, 2020; Sonn et al., 2021).

7. REFERENCES

Carr-Chellman, A., Raney, T., & Campbell, D. (2020). Gem state inequalities: Examining the recent history of Idaho public school funding. *Journal of Education Finance, 45*(4), 407–426. https://www.journalofeducationfinance.com/

Chan, J. (2021). Exploring digital health care: eHealth, mHealth, and librarian opportunities. *Journal of the Medical Library Association, 109*(3), 376–381. https://doi.org/10.5195/jmla.2021.1180

Charles, L. H., & DeFabiis, W. (2021). Closing the transactional distance in an online graduate course through the practice of embedded librarianship. *College & Research Libraries, 82*(3), 370–388. https://doi.org/10.5860/crl.82.3.370

Edwards, M., Kumar, S., & Ochoa, M. (2010). Assessing the value of embedded librarians in an online graduate educational technology course. *Public Services Quarterly, 6*(2-3), 271-291. https://doi.org/10.1080/15228959.2010.497447

Fields, A. (2020). Embedding librarians in online tertiary classrooms: A new model for learner support. *British Journal of Educational Technology, 51*(4), 1373–1385. https://doi.org/10.1111/bjet.12892

Franzen, S. R., & Sharkey, J. (2021). Impact of embedded librarianship on undergraduate nursing students' information skills. *Journal of the Medical Library Association: JMLA, 109*(2), 311–316. https://doi.org/10.5195/jmla.2021.913

Garcia-Castelan, R. M. G., Gonzalez-Nucamendi, A., Robledo-Rella, V., Neri, L., & Noguez, J. (2021). Face-to-face vs. online learning in engineering courses. *Proceedings - Frontiers in Education Conference, FIE, 2021-October.* https://doi.org/10.1109/FIE49875.2021.9637177

Guillot L, Stahr B, & Meeker BJ. (2010). Nursing faculty collaborate with embedded librarians to serve online graduate students in a consortium setting. *Journal of Library & Information Services in Distance Learning, 4*(1/2), 53–62. https://doi.org/10.1080/15332901003666951

Martin, F., Gezer, T., & Wang, C. (2019). Educators' perceptions of student digital citizenship practices. *Computers in the Schools, 36*(4), 238–254. https://doi.org/10.1080/07380569.2019.1674621

Matteson, A. (2020). Building instructional and resource options during COVID. *Teacher Librarian, 48*(2), 4–27. https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib&db=tfh&AN=148603696&site=eds-live&scope=site

Menard, L., & Misquith, C. (2021). Providing real-time resources in support of LGBTQ+ and HIV+ populations as information experts on the ECHO hub team: A case report. *Journal of the Medical Library Association, 109*(4), 631–636. https://doi.org/10.5195/jmla.2021.1262

Mendoza, D. F., & Kerl, E. (2021). Students perceived benefits of embedded online peer tutors. *Learning Assistance Review (TLAR), 26*(1), 53–73. http://www.nclca.org/tlar

Nowicki, J. M. (2020). Disaster recovery: COVID-19 pandemic intensifies disaster recovery challenges for k-12 schools. *GAO Reports*, 1–14. https://www.gao.gov/products/gao-21-62r

O'Brien, C. (2020). The COVID-19 Rollercoaster. *The Learning Assistance Review, 25*(2), 49-59. https://link.gale.com/apps/doc/A637465134/AONE?u=anon~24116e42&sid=googleScholar&xid=078a7f7a

Petrina, S. (2004). Sidney Pressey and the automation of education, 1924-1934. *Technology and Culture, 45*(2), 305–330. https://doi.org/10.1353/tech.2004.0085

Rogers, E. M. (2010). *Diffusion of innovations* (5th ed.). Simon and Schuster.

Shumaker, D. (2014). The Embedded librarian innovative strategies for taking knowledge where it's needed. Information Today. https://books.infotoday.com/books/Embedded- Librarian/Chapter-1.pdf

Silva, P. (2015). Davis' Technology Acceptance Model (TAM) (1989). In M. Al-Suqri, & A. Al-Aufi (Eds.), *Information Seeking Behavior and Technology Adoption: Theories and Trends* (pp. 205-219). IGI Global. https://doi.org/10.4018/978-1-4666-8156-9.ch013

Sonn, I. K., Du Plessis, M., Jansen Van Vuuren, C. D., Marais, J., Wagener, E., & Roman, N. V. (2021). Achievements and challenges for higher education during the COVID-19 pandemic: A rapid review of media in Africa. *International Journal of Environmental Research and Public Health, 18*(24). https://doi.org/10.3390/ijerph182412888

Spangler, S. (2019). Integrating information literacy in IT courses: Information technology students' perceptions of embedded librarians. *Online Journal of Applied Knowledge Management, 7*(2), 29–40. https://doi.org/10.36965/ojakm.2019.7(2)29-40

Spangler, S. C. (2020). IT graduate students' perceptions of embedded librarians. *Issues in Information Systems, 21*(4). https://doi.org/10.48009/4_iis_2020_73-80

Spangler, S.C., Casper, D. & Stanfield, D. (2020). Online students' perceptions of embedded librarians. *Issues in Information Systems. 21*(1), 167-176. https://doi.org/10.48009/1_iis_2020_167-176

Spencer, D., & Temple, T. (2021). Examining students' online course perceptions and comparing student performance outcomes in online and face-to-face classrooms. *Online Learning Journal, 25*(2), 233–261. https://doi.org/10.24059/olj.v25i2.2227

Steele, J. E. (2021). The role of the academic librarian in online courses: A case study. *The Journal of Academic Librarianship, 47*(5). https://doi.org/10.1016/j.acalib.2021.102384

Tumbleson, B. E. (2016). Collaborating in research: Embedded librarianship in the learning management system. *Reference Librarian, 57*(3), 224–234. https://doi.org/10.1080/02763877.2015.11

34376

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186-204. http://www.jstor.org/stable/2634758

Wu, L., Betts, V. T., Jacob, S., Nollan, R., & Norris, T. (2013). Making meaningful connections: evaluating an embedded librarian pilot project to improve nursing scholarly writing. Journal of the Medical Library Association: *JMLA, 101*(4), 323–326. https://doi.org/10.3163/1536-5050.101.4.016Allen, D. B. (2019). The Forgotten Brown case: Briggs v. Elliott and its legacy in South Carolina. *Peabody Journal of Education, 94*(4), 442–467. https://doi.org/10.1080/0161956X.2019.1648954

An Industry Survey of Analytics Spreadsheet Tools Adoption: Microsoft Excel vs Google Sheets

Carl M Rebman, Jr.
carlr@sandiego.edu
Knauss School of Business
Department of Supply Chain, Operations, and Information Systems
University of San Diego
San Diego, CA 92110, USA

Queen E. Booker
queen.booker@metrostate.edu
Department of Management Information Systems
College of Management
Metropolitan State University
Minneapolis MN. 55403, USA

Hayden Wimmer
hwimmer@georgiasouthern.edu
Department of Information Technology
Georgia Southern University
Statesboro, GA 30460, USA

Steve Levkoff
slevkoff@sandiego.edu
Knauss School of Business
Department of Economics
University of San Diego
San Diego, CA 92110, USA

Mark McMurtrey
markmc@uca.edu
School of Business
University of Central Arkansas
Conway AR 72035, USA

Loreen Marie Powell
lpowell@bloomu.edu
Department of Information Technology, Analytics, and Business Education
Commonwealth University of PA, Bloomsburg
Bloomsburg, PA 17815, USA

Abstract

Spreadsheets have long played a key role in business decisions and operations. The use and adoption of data analytics has substantially increased over the last few years and amplified this role. Spreadsheets are often a first tool for data analytics as such applications provide ease of calculation with basic statistics and chart development. For much of the last two decades universities have provided training in Microsoft Excel because that was what companies used and demanded. Since mid-2020, there has been an increase in use of Google Sheets causing some faculty to believe that MS Excel should be replaced. Faculty should always be aware of current and future employer demands and ensure programs meet the expectations of both employers and recent graduates. This study reviews business job postings seeking employees with two or fewer years of work experience between 2019 and 2021 and examines demand for spreadsheet application experience. Results overwhelmingly indicate that Microsoft Excel still is the most required spreadsheet application by employers. Faculty should pause before changing MS Excel training or removing certifications.

Keywords: Curriculum, Spreadsheets, Skills, Position Analysis, Business Education, MS Excel, Google Sheets

1. INTRODUCTION

Graduate employability is a concern for all academic programs as graduate employability has become a measure of program health for program investment at many colleges (AACU, 2021). For business, tools knowledge and expertise are prominent in job descriptions. One such tool that is popular in job descriptions is spreadsheet and concomitant skills such as data analysis and visualization. With an increase in remote work, collaboration is also becoming a more common skill.

Lawson et al. (2009) reported electronic spreadsheet usage in business had been around 25 years. Since their study was published in 2009, it is feasible to extend the time to almost 38 years. According to Melissa James from GreyCampus (2015), spreadsheets are important to business as financial accounting tools, tracing product sales, accessing customer data, analyzing return on investment, and work scheduling. According to Rahman et al. (2020) and Mack et al. (2018) "spreadsheets are everywhere." In addition to business applications, they are used for a wide range of activities such as, but not limited to, managing diets, scientific experiments, real estate inventory, managing grades, managing financial portfolios, and managing real estate inventory. Thus, having good spreadsheet skills is beneficial beyond just knowing how to use them for their business applications.

Higher education business programs play a role in developing students' spreadsheet skills. A recurring question when performing curriculum review is whether faculty are teaching the right skills and tools for employability. To help shape curriculum, it is not uncommon for faculty to review recent position descriptions, survey employers and recent graduates, and review other external sources to ensure that the curriculum provided meets the current and future expectations for both students and employers.

Spreadsheet tools and concomitant skills are an area of interest, especially since spreadsheet dominance has changed through the years. For example, Lotus 1-2-3 replaced VisiCalc, Microsoft Excel replaced Lotus 1-2-3, and now Google Sheets is emerging as a potential tool to replace Microsoft Excel. Understanding how spreadsheet interest is changing in the marketplace helps faculty know when or if a switch in tool emphasis in the curriculum is necessary to ensure graduate employability.

This study examines position descriptions posted on indeed.com, a popular online job posting site. The positions examined were posted between January 1, 2019, through December 31, 2021. The descriptions were limited to business program titles requiring a bachelor's degree, and two years or less of work experience. The position analysis focuses on requirements for spreadsheet knowledge, skills, and abilities such as data analysis and visualization, and on collaboration, a popular skill requirement for remote workers. The remainder of this paper is structured as follows:

The literature review includes a discussion and comparison of Excel and Google Sheets. It also includes information regarding using position descriptions to examine knowledge, skills, and abilities. The literature review is followed by the research methodology and study, results, and then conclusions and next steps.

2. LITERATURE REVIEW

Spreadsheets have existed for several decades. Many years before the first electronic spreadsheet program, LANPAR (LANguage for Programming Arrays at Random, Wikipedia, 2022), businesses used paper-based ledger systems to manage their financial operations. According to Dan Power of DSS Resources.com (2004), "in the realm of accounting jargon, a 'spread sheet' or spreadsheet was and is a large sheet of paper with columns and rows that organizes data about transactions for a businessperson to examine, it shows all the costs, income, and taxes for a manager to examine when making a decision." Accountants primarily used these ledgers to handle budgeting, accounts receivables, and payroll. Accountants found spreadsheets were important as they allowed for better organization of data which enhanced decision making.

Spreadsheets

Since 1969 there have been many spreadsheet applications. Some programs got their start as in-house programs such as Autoplot (General Electric), APLDOT (US Railway Association), while others obtained more commercial success and adoption such as VisiCalc, SuperCalc, Lotus 1-2-3, Microsoft Excel, and Google Sheets. It seemed as if the spreadsheet software world was like that of worldwide boxing champions with each newcomer knocking off the reigning king.

Research and practice have demonstrated that spreadsheet software can be robust and versatile. For example, Grossman et al. (2007) investigated 18 different cases in the areas of application software development, executive information systems, financial risk management, sales and marketing business processes, business operations, and complex analytics. They found that spreadsheets could be more strongly protected as Excel contains tools to keep the developer's source code protected. Large complex spreadsheets can require programming skills similar to a Fourth Generation and Rapid Development Language operating in an Integrated Development Environment.

Reschenhofer and Matthes (2015) discussed how spreadsheets have capabilities like formulae and macros to support complex calculations or automate processes, and spreadsheets have become an indispensable as a comprehensive medium for data management and analysis. Frownfelter-Lohrke (2017) reported that if companies did not create good spreadsheets or conduct thorough spreadsheet analysis then their businesses could suffer from loss of profit or market share.

Microsoft Excel vs Google Sheets

The exact current number of spreadsheet users is debatable and still by all counts considered to be large. Statista (2019) did a study in Finland in 2019 and found that a little under 50% of their population used spreadsheet software and it was relatively evenly spread over different age demographics (Figure 1).

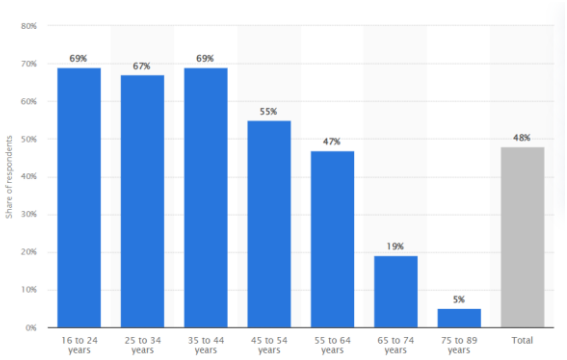


Figure 1: Share of people who use spreadsheet software in Finland, 2019

A few others have tried to determine the actual number of users for spreadsheet software programs. For example, in 2017 Simon Cocking from Irish Tech News wrote about seven reasons that a half a billion people still used Excel. Joseph (2021) estimates that there are 750 million to 1.2 billion Microsoft Excel users and over 2 billion Google Suite users world-wide. Elizabeth Gration (2022) reported that 1 in 8 people use Microsoft although her numbered quote included the entire Office Suite.

In fact, one of the challenges with determining actual Excel and Google Sheets numbers stems from difficulties determining which exact product in the Microsoft Office Suite or Google Suite is being used. Hjalmar Gislason (2018) discusses how it is hard to differentiate between paying and free accounts as well as the inclusion of GMAIL in the Google Suite. To address this, Gisalsson applied a 2/3 estimate reduction in numbers when comparing MS Excel (800 million) and Google Sheets (160 million).

Gilasson (2018) provides an enlightening description of Microsoft Excel and Google Sheets users and is visually summarized in Figure 2. He talks about how more young people and young companies use Google Sheets while older and more mature companies use Microsoft Excel. He reports that Excel is more sophisticated and

preferred by 'pros' where Google sheets is more sluggish but good for quick and simple items. He also did note that Google sheets works better for collaboration and that even the pros were found to 'sneak' using Google sheets from time to time. Nina Semczuk (2020) provides support for Gilasson's claims by confirming the benefits of collaboration in Google Sheets as well as pointing out how frequently Google Sheets gets updated and works well with real time data.

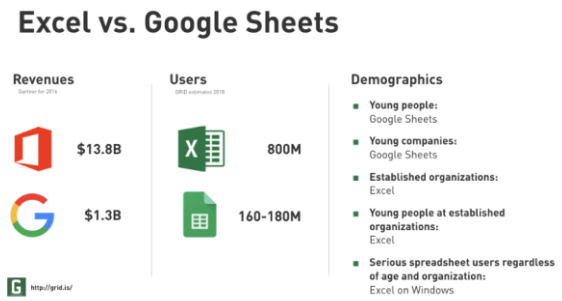


Figure 2: Comparison between Excel and Google Sheets (Source: https://grid.is/)

Perhaps examining the technical differences between the two programs might shed some light on use and adoption. Appendix 1 presents a comparison of the technical differences taken from three sources: Simplilearn, EDUCBA, and WallStreetMojo.com.

All three sources provide the same pros and cons of the two programs. Their analysis reinforces previous research that Excel is more robust by being able to handle more cells and data. Conversely, Google Sheets is better for sharing, collaboration, multi-user, and for updating data. Excel tends to be faster and much better at data visualization, especially with the newer versions of Excel that contain PowerBI integrations and connect to Tableau seamlessly. PowerBI and Tableau are products that allow for easier generations of complex data visualizations.

Newer versions of MS Excel, such as Office 365, contain the Microsoft One Drive feature which was designed to address collaboration and sharing issues. One Drive still has its limitations as it requires more setup than Google Sheets. By comparison, Google Sheets is considered simpler and is continuously updated in real-time as data is entered. For example, Google Sheets automatically updates and saves when users engage with the file, when the file is shared, and it always tracks changes.

Investintech.com (2022) conducted personal interviews with 27 spreadsheet experts and asked them to predict the future of Excel in Business Intelligence. Appendix 2 presents a few of the quotes from the experts. 25 out of 27 felt that Excel would always be important, at the heart of data analytics and highly used and valued.

The main differences noted between Microsoft Excel and Google Sheets were cost, maximum cells per sheet, the robustness, automatic file saving, large datasets, the ability to collaborate effectively, and certification.

In addition to research studies on specific spreadsheet software skills required by employers, our literature review found that studies on spreadsheets focused on two other areas: how to 'teach' spreadsheet software to students in specific domains (e.g., economics, statistics), and the value of certification in spreadsheet skills. Most of the articles were focused on Microsoft Excel, although there were some older articles on VisiCalc and Lotus 1-2-3. Those articles were not included because the software is retired. Articles on Google Sheets are starting to appear and those have focused on how to use and introduce Google sheets either in general terms or towards a specific domain. There was little in term of direct comparison between MS Excel and Google Sheets

Formby et al. (2017) found that 94% agreed or strongly agreed that being able to create spreadsheets, charts and analyze data were especially important skills and that data analytics was also an important skill for students. They reviewed job postings for 5 different states and found that on average approximately 50% of employers had a strong spreadsheet knowledge requirement, specifically Microsoft Excel.

Treadwell et al. (2013) found that industry demands were strong for spreadsheet skills. Their study sought to determine important spreadsheet skills and then test students on Microsoft Excel. They felt the students' performance was declining and that could cause students to have difficulties finding jobs after graduation.

Coleman and Blankenship (2017) attempted to determine the linkage between academia and businesses. They surveyed faculty and area businesses and asked about 18 different MS Excel and MS Access attributes and found many synergies and alignments between what was taught and what was needed by employers. Their study reaffirmed that analytical skills and specifically spreadsheet skills are important to be taught, and to be mastered by students.

Raglan and Ramachandran (2014) investigated student's perception of their MS Excel knowledge and perceived readiness for accounting positions. Their study was helpful in aligning the skills perceived as necessary and important by new hires with those of current students. They also found that Excel is the most common analytical tool used by accountants, that public accounting firms are emphasizing an interest in recruiting accounting students with proficiency in Excel, and they want to hire new staff who know MS Excel.

Lee et al. (2018) reviewed job postings and surveyed accountants. They found that over 50 percent of postings required spreadsheet skills, specifically MS Excel. They found that most accountants felt that spreadsheet skills were very important. Their study even identified specific skills in MS Excel such as 'Lookup and Pivot Table', 'Data entry, format, and calculations', and 'Logic' as necessary for new hires. Their results further indicated that "educational institutions should place emphasis on spreadsheet training for accounting students, if they have not already done so" (Lee et al., 2018).

Many of the studies on how to teach spreadsheets to students are predominately focused on Microsoft Excel. Barreto (2015) provided an interesting review of Microsoft Excel pedagogical history in Economics. They found spreadsheets improve learning outcomes across the Economics curriculum and are increasingly being used. His article was relatively unique in that it did not focus on a specific application or example which is more common in the literature. Barreto (2022) also published a paper using Microsoft Excel to teach microeconomics.

In addition to Economics, studies were found for using Microsoft Excel to teach in other domains. For example, Divisi et al. (2017) presented a paper on statistics and how Microsoft Excel has functions to help with learning statistical processes. Chaamwe and Shumba (2016) wrote a paper using Excel to teach statistics in an e-learning environment. Khan (2007) wrote about using Excel to teach physics. Al Rawahi et al. (2006) demonstrated Excel in math. Munisamy (2009) showed how Excel could be of benefit to operations research. Willis (2016) implemented Excel in accounting information systems. Mangeiro et al. (2010) provided an example of using Excel in a financial management class. Conant and Chaille (2022) created an excel-based classroom exercise to demonstrate an asset depreciation method comparison.

From the certification perspective, Clairborne (2017) did a comprehensive review of employers and employer desired certifications in prospective employees. They discovered that spreadsheet skills were viewed very positively, and the Microsoft Office Specialist (MOS) Exam in Excel was considered the most valuable out of all the Microsoft Office suite products. Rebman et al. (2019) found that certifications help prepare students to compete in competitive job markets and showcase their marketability while they are still in school.

Most of the studies involving Google Sheets were focused on introducing a new alternative to Excel or how to apply Google Sheets to classroom instruction. For example, Parra et al. (2021) created a teaching case to teach cloud-based Google Sheets using Shippy Express. The goal of the teaching case was to have students use Google Sheets to develop summaries of their financial transactions to help them make decisions. Blair and Mahoney (2022) present a method for using Google Sheets to provide a step-by-step system for creating graphs for research designs and clinical applications. Ovezmyradov (2020) illustrated a way to use Google Sheets to teach the classic supply chain beer game.

There were a few studies that conducted comparison studies. Lawson et al. (2009) to 1600 subjects and put them into two groups of spreadsheet level and found that people with spreadsheet experiences and training did much better. They also found that corporate culture played a role and those who valued training had more advanced users. It was unclear if other companies relied on new hires to have spreadsheet experience. This was one of a few studies by the Spreadsheet Engineering Research Project (http://mba.tuck.dartmouth.edu/spreadsheet/) and focused on Microsoft Excel.

However, the lack of research on spreadsheet usage has not gone unnoticed. Grossman (2007) commented on the lack of spreadsheet research in comparison to research on other programming languages despite its wide use and impact on business. Today there are still gaps in spreadsheet research, specifically as it relates to a determination on the more preferred spreadsheet software of business, and trends in spreadsheet skills, knowledge, and abilities. There is a fair amount of Reddit chat boards or simple blog reviews and not so much in peer-reviewed research. This study seeks to address this gap in research.

Position Description Analysis and Curriculum Development

Current literature has examined methodologies faculty can use to determine if their programs are adequately preparing students for their current workforce environment. Downey et al. (2008) mentioned that examining job postings was an excellent method to determine the in-demand industry skill and tools. This approach was supported in Harper (2012), McArthur et al. (2017), and Munmun et al. (2022), who similarly found that examining job postings can successfully enhance the understanding of the industry's expectation for particular skills and tools and as well as the employment environment. Harper (2012) and McArthur et al. (2017) discussed the importance of reviewing position announcements in a specific area where the industry demand is emerging. Overall, job advertisements can successfully show the hiring trends, labor demand analytics, and specific skill requirements, and can be the most effective and reliable resources to design curriculum within and across courses (Ahsan et al., 2013; Carnevale et al., 2014; Carnevale et al., 2014; Diamond et al., 2014; Frankenfeld, 2017; Hirudayaraj & Baker, 2018; Meyer, 2017; Reeves & Hahn, 2010; Rosén, 2014; Stanton, 2017; Templin and Hirsch, 2013, Wellman, 2010).

Desai & Von Der Embse (2001) and Uğur & Hamit Turan (2019) suggest strategies to develop effective curriculum based on practitioner surveys built on the existing curriculum literature and job postings, one of which focused on concentrations and the other on academic departments and industry to collaborations to understand current and future emerging needs. proposed a "synergistic approach" to employ both integrative and intensive strategies in higher education program development. The authors focus on "assessing industry trend for specific IT". This "assess industry trend" means to find out the promising tools and skills associated with them, then assessing how complex/difficult they are, alignment with student body, and local industry needs to determine what should be included in curriculum.

3. RESEARCH METHODOLOGY AND STUDY

This study seeks to understand the trends in employer perception and value of spreadsheet skills and preference using employer position analysis. Specifically, the study compares employer expectations between Microsoft Excel and Google Sheets. To provide a context, the study also sought to determine the significance of data analysis skills. Given the historical displacement of spreadsheet applications and the emergence of Google Sheets as a competitor in the spreadsheet application market, it is important to study its demand in comparison to the demand for Microsoft Excel in position descriptions and understand how the changing work environment may be impacting employer expectations. Concomitantly, understanding the changing expectations of employers can help faculty in higher education respond to that change to maintain student preparation to perform in the workplace. This study uses position analysis to determine if faculty should consider updating courses to include Google Sheets as part of the future curriculum as well as measuring demand for particular analysis skills that can be performed using spreadsheets.

The literature review demonstrated the link between data analysis skills and spreadsheets. (Formby et al. (2017); Treadwell et al. (2013); Coleman and Blankenship (2017); Raglan and Ramachandran (2014); and Lee et al. (2018)). Specific skills considered in the literature included the ability to create spreadsheets, charts and data analysis. This led to the question of how important these skills are currently. The first research question and hypotheses are:

1. Do employers consider having the ability to create spreadsheets, charts, and analyze data to be an important skill for students and new hires in 2021 as compared to 2019?

H1: The prevalence of general spreadsheets knowledge, skills and abilities for new hires is significantly higher in 2021 than in 2019.

H1a. The prevalence of data analysis knowledge, skills and abilities for new hires is significantly higher in 2021 than in 2019.

H1b. The prevalence of developing charts knowledge, skills and abilities for new hires is significantly higher in 2021 than in 2019.

The literature review also discussed the strengths and weaknesses of Microsoft Excel and Google Sheets as tools (Simplilearn, EDUCBA, and WallStreetMojo.com). These strengths and weaknesses are important based on the expectations of the position.

While it is important to study the overall expectations of spreadsheets knowledge, skills and abilities, specific tool knowledge is also important. Given that Microsoft products have been on the market longer, it may be expected that Microsoft Excel would be listed as a required

or preferred knowledge, skill or ability more often than Google Sheets. Thus, the next research question and hypothesis are:

2. Do employers prefer Microsoft Excel skills over Google Sheets skills?

H2: The prevalence of MS Excel knowledge, skills and analysis is significantly higher than Google Sheets in positions posted.

Since Google Sheets was initially designed as a collaborative tool (Simplilearn, EDUCBA, and WallStreetMojo.com), it may be considered a stronger collaborative tool than Excel. Although collaboration is an expectation in the workplace in general, the increase in remote work during 2020 and 2021 may have created a demand for the ability to collaborate electronically including using tools such as Google Sheets.

Thus, it seems relevant to analyze if there is correlation between the spreadsheet tools and work environment expectations such as collaboration and working remotely. The next research questions and hypotheses studied were:

3. Do employers who list positions as remote prefer Microsoft Excel or Google Sheets?

H3. There is no significant correlation between remote work environments and preference for Microsoft Excel or Google Sheets in posted job positions.

4. Do employers who list collaboration as a required or preferred skill prefer either Microsoft Excel or Google Sheets?

H4. There is a higher correlation between collaboration skills and Google Sheets than for collaboration skills and Microsoft Excel.

Several statistical methods were used to address and provide a response to the four research questions. We analyzed job postings from Indeed.com for the years 2019 to 2021.

Indeed.com is a free service to job seekers, where employers post position openings and applicants can upload a resume, create job alert emails, search for jobs, save them and apply to them directly. (Indeed.com, 2022) Positions were extracted from the indeed.com data using the following requirements:

- Position Keywords: accounting, management, information systems, marketing, sales, operations, supply chain, logistics, economics, human resources
- Degree restriction: Bachelor's required or preferred
- Experience restriction: Two years or less

This extraction resulted in 21,398 non-duplicated position descriptions of which there were 5,970 positions for 2019, 7528 for 2020, and 7900 for 2021. There were 284 remote positions posted in 2019, 2160 in 2020 and 1794 in 2021. Table 1 shows the breakdown in position descriptions that requirements for general spreadsheet, data analytics and collaboration skills.

YEAR	COLLAB-ORATION	DATA ANALYTICS	SPREAD-SHEETS
2019	4047 *(43%)*	1061 *(11%)*	4378 *(46%)*
2020	5502 *(43%)*	1426 *(11%)*	5777 *(46%)*
2021	5672 *(41%)*	1646 *(12%)*	6402 *(47%)*
TOTAL	15221	4133	16557

Table 1: General Skills Requirements 2019-2021

The positions were analyzed using text extraction to identify the keywords relevant to the study. The keywords used for the extraction tool were collabor* for collaboration, remote to identify remote positions, Microsoft+Excel, Google+Sheets, data+analytics, visualization, charts and or graph for charts; pivot+tables, statistics and/or data+analysis for data analysis; and spreadsheets for spreadsheets-general. If a keyword was found in the position description, the observation was coded with a 1, and 0 otherwise. The final list of variables was YEAR, MS EXCEL, SHEETS, REMOTE, COLLABORATION, and DATA ANALYTICS.

All the variables except YEAR were coded as binary (0,1). The breakdown in positions that specific Microsoft Excel and/or Google is shown in Table 2. Figure 3 presents the most frequently occurring software requirements overall job postings over the years 2019-2022 and illustrates how Excel has held a steady rate while Google Sheets experienced an increase between 2020 and 2021. It is important to note that many people were forced to work remotely in parts of 2020 and 2021.

YEAR	MS EXCEL	GOOGLE SHEETS	BOTH
2019	3587 *(61%)*	1243 *(21%)*	957 *(18%)*
2020	4546 *(57%)*	2008 *(25%)*	1395 *(18%)*
2021	4724 *(51%)*	2899 *(31%)*	1715 *(18%)*
TOTAL	12857	6150	4067

Table 2: Positions mentioning MS Excel, Google Sheets and/or both.

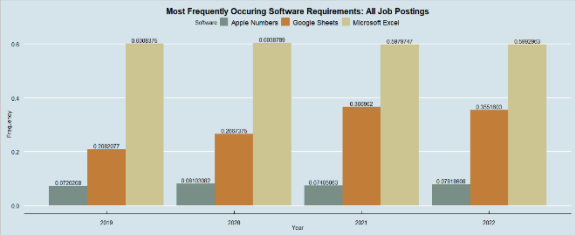


Figure 3: Most Frequently Occurring Software Requirements

The variables were analyzed using parts of the approach in Uğur, & Hamit Turan (2019). Since all the data is binary, the statistical analysis used t-tests for those hypotheses that compared pairs of data and correlation for the hypotheses that focuses on relationships between values.

4. RESULTS

The data variables analyzed in the study were MS EXCEL, GOOGLE SHEETS, SPREADSHEET-GENERAL, REMOTE, CHARTS, COLLABORATION, and DATA ANALYTICS.

A correlation matrix was created for the variables to determine the strength of the relationship between remote work, collaboration and the two tools. The correlation between remote work and Google Sheets was .038 as compared to .014 for remote work and MS Excel indicating that Google Sheets was a slightly more preferred tool than Excel for those working remotely. Similarly, the correlation between collaboration and Google Sheets was .17 as compared to .16 for collaboration and MS Excel, indicating that Google Sheets was a more preferred tool than Excel for those positions stating collaboration as a required or preferred skill – but not by much.

However, when the correlations were performed by year, the correlation between remote work and Google Sheets increased from .002 in 2019 to respectively to .019 in 2021 as compared to .0024 in 2019 to .013 for Excel and remote work indicating that Google Sheets is growing faster as a correlated skill for remote work than Excel. Similar results were found for collaboration. The correlation matrices can be found in Tables 3, 4 and 5.

Overall Correlation Matrix (2019, 2020 and 2021)		
Tool	*Remote Work*	*Collaboration*
Google Sheets	0.038	0.17
Microsoft Excel	0.014	0.16

Table 3: Overall Correlation Matrix

Correlation Matrix for 2019		
Tool	*Remote Work*	*Collaboration*
Google Sheets	0.002	0.12
Microsoft Excel	0.0024	0.11

Table 4: Correlation Matrix for 2019

Correlation Matrix for 2021		
Tool	*Remote Work*	*Collaboration*
Google Sheets	0.019	0.16
Microsoft Excel	0.013	0.13

Table 5: Correlation Matrix for 2020

The t-tests comparing Google Sheets demand to MS Excel demand had a p-value of .000 with means of .3 and .6, respectively. The t-test comparing Google Sheets demand, MS Excel demand, data analytics demand, and spreadsheet-general demand between 2019 and 2021 found all the comparisons to be significant with p-values of .000 with higher means for 2021 than for 2019.

The last t-test compared the 2019 and 2021 demands for general spreadsheet skills, data analysis and charts. The results indicate higher demand in 2021 than in 2019 for each as follows:

- Data analysis had p-value of .000 with means of .6 and .45, respectively
- Charts had p-value of .000 with means of .56 and .31, respectively
- Spreadsheet-general had a p-value of .000 with means of .7 and .56, respectively.

Based on the results, all hypotheses can be accepted. However, for the question of prevalence of MS Excel or Google Sheets, it should be noted that based on the trend analysis,

Google Sheets demand is growing slightly faster than MS Excel as a knowledge, skill or ability and that growth is correlated with growth in remote work.

Further, the results also suggest that faculty should be cognizant of the trends for increasing demand for data analysis skills and spreadsheets skills in general, as well as the increase in demand for collaboration skills and the number of positions that allow for remote work. There should be curriculum consideration of a balance between face-to-face and remote work in programs to ensure students become comfortable with both. Also, faculty should analyze and ensure sufficient collaborative projects that require online and face-to-face interactions to help students build collaboration skills in both environments. Lastly, faculty should examine the depths and applications of data analysis and spreadsheet skills throughout the curriculum to ensure that students are exposed to a variety of skills including but not limited to statistical analyses such as trend, descriptive statistics, and regression; pivot tables; and effective chart and graph techniques.

As to whether there should be an emphasis on Microsoft Excel or Google Sheets, though, the answer is less definitive. Although Microsoft Office 365 was introduced in 2011, it has increased its collaborative tools with Teams and OneDrive, and offers multiple ways to share documents. However, the ability to simultaneously edit documents in Google Suite is probably one of the reasons that demand for Google products is increasing. Further, this analysis was a national view rather than a regional view. Faculty are encouraged to analyze the position descriptions for their regions to determine if the national results hold true for their specific region.

5. LIMITATIONS AND FUTURE RESEARCH

This study provided a US-based lens for specific skills for new university bachelor-level graduates. While a national view is helpful, individual campuses would benefit from an analysis based on the student population it serves. For example, if the student population is largely regional, then a regional perspective is needed. If the student population has a significant percentage of international students, then a global perspective should be considered. This is perhaps the highest limitation of the study. The data is also from one source. One of our next steps is to analyze the data by states or cities and compare regional expectations.

Further, 2020 and 2021 represent the two years with most companies asking workers to work remotely due to a pandemic. This may artificially impact positions' demand for remote work and collaboration. We also intend to examine the 2022 data to determine if the trends found for 2019 through 2021 continued in 2022.

Although the data was validated with three people reviewing a random selection of the observations for algorithm accuracy in extracting results, it is likely there are other keywords that were missed that might suggest a preference for specific data analytics or analysis skills.

Another limitation is the scope of this study. Examining job postings only partially answers the question of which software program should be taught. Companies have cultures to which have their own norms and behaviors and can influence decision making. Surveying employers on their software preference, value, and adoption could also help provide more answers as to what should be taught.

Lastly, previous studies (Bakir et al., 2019; Rebman et al., 2021; Tastle et al., 2017) have discussed Microsoft Excel certification curriculum programs and their impact on producing strong Excel knowledge workers. Future research would be helpful to determine if employers reported the having the same perception and value of MS Excel certified employees. It would be interesting to combine the results of employer preference to employer perceived value of certification and compare it with employer job postings. Such are questions we learned from this study that we plan to address in the future.

6. CONCLUSION

This study was conducted as a response to questions regarding curriculum, specifically whether Microsoft Excel should be taught or replaced with Google Sheets. The results reinforce previous studies in that MS Excel is still preferred over Google Sheets, at least in terms of job postings. The increase of remote jobs indicates that Google sheets should not be disregarded. As companies determine their new normal, there may be increased demand for products that promote and support collaboration the way Google products do.

As technology changes, it is important that faculty analyze regional or local trends to determine if their respective curriculums adequately address knowledge, skills and abilities required by employers. However, there doesn't appear to be

an immediate need to remove MS Excel training and certifications from the curriculum or to replace course work with alternatives such as Google Sheets.

7. REFERENCES

Ahsan, K., Ho, M., and Khan, S. (2013). Recruiting project managers: A comparative analysis of competencies and recruitment signals from job advertisements. *Project Management Journal,* 44(5), 36-54.

Al Rawahi, F. K., Khan, S. A., & Huq, A. (2006). Microsoft Excel in the mathematics classroom: a case study. *METSMaC 2006*.

Bakir, N., Dana, K., & Abdullat, A. (2019). Certifying Business Students in Microsoft Office Specialist Certification Excel Core Exam: Lessons Learned. *Information Systems Education Journal*, *17*(6), 4-11.

Barreto, H. (2015). Why excel?. *The Journal of Economic Education*, *46*(3), 300-309.

Barreto, H. (2021). *Using Excel to Teach Principles of Microeconomics* (No. 2021-01).

Blair, B. J., & Mahoney, P. J. (2022). Creating Single-Subject Research Design Graphs with Google Applications. *Behavior Analysis in Practice*, *15*(1), 295-311.

Carnevale, A. P., Jayasundera, T., & Repnikov, D. (2014). Understanding online job ads data: a technical report. *Georgetown University, McCourt School of Public Policy, Center on Education and the Workforce*.

Chaamwe, N., & Shumba, L. (2016). ICT integrated learning: Using spreadsheets as tools for e-learning, a case of Statistics in Microsoft Excel. *International Journal of Information and Education Technology*, *6*(6), 435-440.

Claiborne, M. (2017). An Examination of Microsoft Office Specialist Certifications and Employability Skills Sought by Georgia Employers. https://etd.auburn.edu/handle/10415/5828 [accessed 24 March 2023]

Cocking, S. (2017). Seven Reasons why Excel is still used by half a billion people worldwide. *Irish Tech News* https://irishtechnews.ie/seven-reasons-why-excel-is-still-used-by-half-a-billion-people-worldwide/ 13 December 2017 [accessed 24 March 2023]

Coleman, P. D., & Blankenship, R. J. (2017). What Spreadsheet and Database Skills Do Business Students Need?. *Journal of Instructional Pedagogies*, 19, 1-8.

Conant, D., & Chaille, S. (2022). Asset depreciation method comparison: An excel-based classroom exercise. *Journal of Education for Business*, 97(5), 351-356.

Desai, M. S., and Von Der Embse, T. (2001). A synergistic strategy for MIS curriculum development: response to rapidly advancing information technology. *College Student Journal,* 35(4), 552-560.

Diamond, K., Pierce, D., Johnson, J., and Ridley, M. (2014). Content analysis of sponsorship sales job postings in the United States. *Graduate Journal of Sport, Exercise, & Physical Education Research* (2), 19-36.

Divisi, D., Di Leonardo, G., Zaccagna, G., & Crisci, R. (2017). Basic statistics with Microsoft Excel: a review. *Journal of thoracic disease*, *9*(6), 1734-1740.

Downey, L. A., Lomas, J., Billings, C., Hansen, K., & Stough, C. (2014). Scholastic success: Fluid intelligence, personality, and emotional intelligence. *Canadian Journal of School Psychology,* 29(1), 40-53.

EDUCBA (2022). "Differences between Excel and Google Sheets" https://www.educba.com/excel-vs-google-sheets/ [accessed 24March 2024]

Formby, S. K., Medlin, D., & Ellington, V. B. (2017). Microsoft Excel®: is it an important job skill for college graduates?. *Information Systems Education Journal*, 15(3), 55-63.

Frankenfeld, C. L. (2017). Trends in employer postings for epidemiology jobs: an analysis of PublicHealthJobs net data from 2003 to 2016. *Annals of Epidemiology,* 27(9), 553-557.

Frownfelter-Lohrke, C. (2017). Teaching good Excel design and skills: A three spreadsheet assignment project. *Journal of Accounting Education*, 39, 68-83.

Gartner, (2020, September). "Gartner Says Worldwide Robotic Process Automation Software Revenue to Reach Nearly $2 Billion in 2021. Retrieved from: https://www.gartner.com/en/newsroom/press-releases/2020-09-21-gartner-says-worldwide-robotic-process-automation-software-revenue-to-reach-nearly-2-billion-in-2021 [accessed 24 March 2023]

Gislason, Hjalmar (2018). Excel vs. Google Sheets usage — nature and numbers. *Medium https://medium.grid.is/excel-vs-google-sheets-usage-nature-and-numbers-*

9dfa5d1cadbd 28 April 2018 [accessed 24 March 2023]

Gration, E. (2022). Microsoft Excel Statistics: Spreadsheets by Numbers. *Micro Biz Mag* https://www.microbizmag.co.uk/microsoft-excel-statistics/# 21 April 22 [access 24 March 2023]

Grossman, T. A., Mehrotra, V., & Özlük, Ö. (2007). Lessons from mission-critical spreadsheets. *Communications of the Association for Information Systems*, 20(1), 60, 1009-1042.

Harper, S. R. (2012). Black male student success in higher education: A report from the National Black Male College Achievement Study. *Center for the Study of Race and Equity in Education, University of Pennsylvania Graduate School of Education*.

Hirudayaraj, M., and Baker, R. (2018). HRD competencies: analysis of employer expectations from online job postings. *European Journal of Training and Development*, 42(9), 577-596.

INDEED.COM (2022). https://www.indeed.com/?from=gnav-employer--post-press--jobseeker [accessed 24 March 2023]

INVESTINTECH (2022). 27 Microsoft Excel Experts Predict the Future of Excel in Business Intelligence. *Investintech.com* https://www.investintech.com/resources/blog/archives/5718-experts-predict-the-future-of-excel-in-business-intelligence.html [accessed 24 March 2023]

James, M. (2015). Why Microsoft Excel is Important for Business. *GREY CAMPUS* https://www.greycampus.com/blog/workplace-tools/why-microsoft-excel-is-important-for-business-organizations 23 November 2015 [accessed 24 March 2023]

Joseph, N. (2021). Number of Google Sheets and Microsoft Excel. *WONDER* https://askwonder.com/research/number-google-sheets-users-worldwide-eoskdoxav 25 February 2021. [accessed 24 March 2023]

Khan, S. A. (2007, March). Microsoft Excel in the physics classroom. In *The Third Annual Conference for Middle East Teachers of Mathematics, Science and Computing, The Petroleum Institute, Abu Dhabi,* 171-175.

Lawson, B. R., Baker, K. R., Powell, S. G., & Foster-Johnson, L. (2009). A comparison of spreadsheet users with different levels of experience. *Omega*, 37(3), 579-590.

Lee, C. B. P., Tang, H., Sam, K. M., & Xiong, G. (2018). Spreadsheet proficiency: Which spreadsheet skills are important?. *J. Inf. Technol. Manag.*, 29(3), 35-44.

Lee, S., Koh, S., Yen, D., and Tang, H. (2002). Perception gaps between IS academics and IS practitioners: An exploratory study. *Information and Management, 40(1), 51 – 61.*

Mack, K., Lee, J., Chang, K., Karahalios, K., & Parameswaran, A. (2018, April). Characterizing scalability issues in spreadsheet software using online forums. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CS04, 1-9.

Mangiero, G. A., Manley, J., & Mollica, J. T. (2010). Improving pedagogy through the use of dynamic Excel presentations in financial management courses. *American Journal of Business Education (AJBE)*, 3(1), 91-106.

McArthur, E., Kubacki, K., Pang, B., and Alcaraz, C. (2017). The employers' view of "work-ready" graduates: A study of advertisements for marketing jobs in Australia. *Journal of Marketing Education*, 39(2), 82-93.

Munisamy, S. (2009). A Spreadsheet-Based Approach for Operations Research Teaching. *International Education Studies*, 2(3), 82-88.

Munmun, M., & Booker, Q. (2022) Industry RPA Demands and Potential Impacts for MIS and Related Higher Education Programs, AMCIS 2022 (*forthcoming)*

Ovezmyradov, B., Meuthia, Y., & Kurata, H (2016). Modeling and teaching inventory management in multiplayer supply chain competition game using Excel and Google Sheets. https://www.researchgate.net/publication/338178065 [accessed 24 March 2023]

Parra, F., Jacobs, A., & Trevino, L. L. (2021). Shippy Express: Augmenting accounting education with Google Sheets. *Journal of Accounting Education*, 56(100740), 1-20.

Powell, S. G., Baker, K. R., & Lawson, B. (2008). A critical review of the literature on spreadsheet errors. *Decision Support Systems*, 46(1), 128-138.

Power, D. J. (2004). A Brief History of Spreadsheets. DSSResources.COM, http://dssresources.com/history/sshistory.html, Version 3.6, 30 August 2004 [accessed 24 March 2023]

Ragland, L., & Ramachandran, U. (2014). Towards an understanding of excel functional skills needed for a career in public accounting: Perceptions from public accountants and accounting students. *Journal of Accounting Education*, 32(2), 113-129.

Rahman, S., Mack, K., Bendre, M., Zhang, R., Karahalios, K., & Parameswaran, A. (2020, June). Benchmarking spreadsheet systems. In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data,* 1589-1599.

Rebman, C., White, G., Wimmer, H., Powell, L. M., & Booker, Q. E. (2021). Pandemic Shift: Impact of COVID-19 on IS/Microsoft Office Specialist Excel Certification Exam Classes--Remote Testing and Lessons Learned. *Information Systems Education Journal*, 19(6), 4-12.

Reeves, R. K., and Hahn, T. B. (2010). Job advertisements for recent graduates: Advising, curriculum, and job-seeking implications. *Journal of education for library and information science*, 51(2), 103-119.

Reschenhofer, T., & Matthes, F. (2015, June). An empirical study on spreadsheet shortcomings from an information systems perspective. In *International Conference on Business Information Systems,* Proceedings 18, 50-61. Springer International Publishing.

Rosén, M. E. (2014). From ad-man to digital manager: Professionalization through Swedish job advertisements 1960-2010. *Journal of Communication Management*, 18(1), 16-39.

Semczuk, N. (2020). 4 Times You Should Really Be Using Google Sheets Instead of Excel. *The Muse* https://www.themuse.com/advice/4-times-you-should-really-be-using-google-sheets-instead-of-excel [accessed 24 March 2023]

Simplilearn (2022). Excel vs Google Sheets: What you need to know?. *Simplilearn.com* https://www.simplilearn.com/tutorials/excel-tutorial/google-sheets-vs-excel 5 July 2022 [accessed 24 March 2023]

Stanton, R. (2017). Do technical/professional writing (TPW) programs offer what students need for their start in the workplace? A comparison of requirements in program curricula and job ads in industry. *Technical Communication,* 64(3), 223-236.

Statista (2019) https://www.statista.com/statistics/554452/spreadsheet-software-usage-by-age-group/ 27 June 2022 [accessed 24 March 2023]

Tastle, W., Mead, C., Rebman, C., Marks, S., & Phillips, K. (2017). Building Excel Expertise: A Guide in Best Practices. In *Proceedings of the EDSIG Conference*, 3(4326), 1-8.

Templin, T., and Hirsch, L. (2013). Do online job ads predict hiring. *New York: New York City Labor Market Information Services.*

Thakar, M. (2022). Excel vs. Google Sheets. *WallStreetMojo.com* https://www.wallstreetmojo.com/excel-vs-google-sheets/ [accessed 24 March 2023]

Treadwell, G., Estep, M., Smith, K. D., & Merritt, K. L. (2013). Spreadsheet proficiency in business school students: A preliminary study of student job preparedness. *Association of Business Information Systems (ABIS) Refereed Proceedings*, 87-100.

Uğur, N. G., and Hamit Turan, A. (2019). Critical professional skills of MIS graduates: Practitioner vs. Academician perspectives. *Journal of Education for Business,* 94(4), 251-258.

Webb, G. K. (2006). The market for IS and MIS skills and knowledge: analysis of on-line job postings. *Issues in Information Systems,* 7(1) 25-258.

Weber, L. (2015). The Key to a Good-Paying Job Is… Microsoft Excel. *Wall Street Journal*. https://pbdd.org/wp-content/uploads/2016/07/The-Key-to-a-Good-Paying-Job-Is_Microsoft-Excel_-At-Work-WSJ.pdf [accessed 24 March 2023]

Wellman, N. (2010). The employability attributes required of new marketing graduates. *Marketing Intelligence & Planning,* 28(7) 908-930.

Willis, V. F. (2016). A model for teaching technology: Using Excel in an accounting information systems course. *Journal of Accounting Education*, *36*, 87-99.

Appendix 1. Microsoft Excel and Google Sheets Comparison

		Excel	Google Sheets
	Software Developed By	**MS Excel, 1987**	**Google Sheets, 2006**
	Certification	**YES, MOS**	**NO, discontinued**
	**Languages**	91 Languages	83 Languages
	Price	Expensive, Office 365 online $8.25 a month	Cheap. Free. Business subscriptions is $5 a month
	**Large Data Sets**	Ideal for large data sets	As data size increases, Google slows down. More suitable for small datasets.
	**Collaboration**	No default to share file users must set up connection	Google Sheet default set up for collaboration
	**Chatting**	Not Available	Available in sidebar
	**Tracking**	Available, but manual	Automatically available
	**Online/offline**	Can be used on &offline	Can be used on and offline
	**Cell Quantum**	Around 17 billion cells available	Maximum 2 million cells available
	**Number of Functions**	Supports 400 functions	supports limited functionality
	**Tool for Statistical Analysis & Visualization**	Superior for data visualization since many formulas are built-in in Microsoft Excel.	Only simple chats, Gantt or a flow chart must be done manually
	**Seamless and easy to use**	Excel is easy to use, needs manual file save.	File automatically saved to Google Drive
	**Usage of macros**	Has built in macros, custom macros, VBA or find premade one	Google Sheets has just added macro functionality

		Excel	Google Sheets
	Software Developed By	**MS Excel, 1987**	**Google Sheets, 2006**
	Certification	**YES, MOS**	**NO, discontinued**
	Linkage/sync with external data	Data needs to be imported	Data can be imported directly from Internet

Appendix 2 Sample list of Experts on Predicting Spreadsheet software (Invenstintech.com)

There is a running joke in BI communities.
"What is the most used feature in any business intelligence solution?"
It is the Export to Excel button.

Excel continues to be the #1 platform when it comes to analyzing data, finding information, preparing charts and presenting them to decision makers. In that sense, I see Excel playing a strong role in BI workflows in future.
/* -Purna Duggirala */

Excel will be right where it always has been – at the center, loved by its users, disliked by IT and the target of endless attempts to replace it by third party BI software vendors.
/*- Chris Webb*/

BI makers understand that Excel is here to stay, it is versatile and works excellently with the web and BI systems.
/* – Tom Urtis*/

The addition and integration of the new Power BI tools within Excel really opens up the possibilities for any type of analyst or everyday user to get involved in Business Intelligence.
/* – Jon Acampora*/

Not everyone works with millions of rows of data. What matters to everyone, anywhere, is what they can do with the data.
/* – Andrew Engwirda */

The user will ultimately define how Excel fits into the BI workflows. It could happen at any point in the BI process.
/* – Jordan Goldmeier*/

Onliners versus On-grounders in Computer and Information Systems courses in Higher Education: A Two-Step Cluster Analysis

Alan Peslak
arp14@psu.edu
Information Sciences and Technology
Penn State University
Dunmore, PA 18512 USA

Lisa Kovalchick
kovalchick@pennwest.edu
Computer Science and Information Systems
Pennsylvania Western University
California, PA 15419 USA

Wenli Wang
wangw@rmu.edu

Paul Kovacs
kovacs@rmu.edu

Computer and Information Systems
Robert Morris University
Moon Township, PA 15108 USA

Abstract

Are students who prefer online education different from those who prefer on-ground education, and how? This is an important question because educational institutions need to better understand student segmentations. This research examined 251 survey responses from students enrolled in Computer Information Systems courses at three universities over five years (2016-2021) and reviewed student attitudes, perceived skills, and their sociological characteristics. Through two-step cluster analysis unsupervised machine learning, two distinct clusters of students emerged, namely Onliners and On-grounders. The top nine out of the eleven group characteristics for Onliners are: select more online courses, regard online instruction as effective, work better without supervision, rely less on classroom interactions in learning, value convenience, can prioritize, are better organized, better prepared, and older. By understanding these group characteristics, educational institutions can make better decisions in policy making, resources allocation, and student recruitment/retention.

Keywords: Online education, on-ground education, cluster analysis, machine learning, persona, cluster

1. INTRODUCTION

As the onset of the COVID-19 pandemic brought about an abrupt transition from traditional on-ground education to online learning, online education has become a focal point of research for educators. Although online education existed pre-pandemic, the pandemic has made it much more prevalent. Now with the shift to more relaxed COVID-19 restrictions and, hopefully, the end of the pandemic, decision-makers in higher education face challenges in making the right decisions concerning online versus on-ground education such as policy making, resource allocation, and student recruitment/retention.

These decision-making challenges concerning online versus on-ground education have resemblance to those faced by CEOs concerning remote versus in-office work. Some firms such as Yelp, AirBnB, 3M, Lyft, and Spotify have gone fully-remote (Lufkin, 2022). PayPal posts both "fully remote" and "opt for remote" jobs. SAP allows employees to choose from remote, in-office, or hybrid work (Smith, 2022). JPMorgan Chase's CEO, Jamie Dimon, on the contrary, has a long-held preference of in-office work (Shevlin, 2022).

Despite some CEO's preference, many firms have based their decisions of employee work locations on meeting employees' expectations. Microsoft surveyed over 31,000 employees in 31 countries in 2022, and 52% of them were willing to switch to fully-remote or hybrid jobs (Microsoft, 2022). Employees are now more likely to prioritize their health and wellbeing over work; this is especially true of employees who are parents and/or women (Microsoft, 2022).

Although in-office work may help strengthen culture, improve collaboration, and reinforce purpose (Markman, 2021) and remote work may reduce costs and offer flexibility, the long-term impact of remote work is inconclusive. Similarly, Peslak, Kovalchick, Wang, and Kovacs (2021) showed mixed-results pertaining to online education. Just as employees in different clusters have varied expectations and performances, students in different clusters have varied attitudes and learning effectiveness (Peslak, Kovalchick, Wang, & Kovacs, 2021). For instance, Bishop (2022) studied employees at different age groups and concluded that 81% of under-35-year-olds fear loneliness from long-term homeworking. Similarly, do younger students fear more adverse effects of online education?

There is limited literature that has examined student clusters and investigated whether there are distinct characteristics of students' preferring online education versus on-ground education. This research aims to fill in the knowledge gap and answer the following research questions.

RQ1: What are the subgroup models that emerge from a multi-year, multi-university student learning preferences survey?

RQ2: What are the characteristics of the student subgroups that emerge from a multi-year, multi-university student learning preferences survey?

The first research question was modeled after Stewart, Miller, Audo, and Stewart (2012) who used cluster analysis to identify patterns in student responses. The second research question aims to provide more insights of student segmentation.

This research employs cluster analysis in order to better understand online versus on-ground education and the groups of individuals that may have significantly different views toward these educational modes. Cluster analysis allows us to identify distinct student subgroups and their characteristics. Providing understanding on student segmentation, this research provides insights into the types of students who prefer online versus on-ground education. It provides decision support regarding policy making, resource allocation, and how to better market to, recruit, and retain students within potentially distinct subgroups.

2. LITERATURE REVIEW

Online versus On-ground Education

In 2020, as the arrival of the COVID-19 pandemic closed university campuses, a record number of students were forced into online classes. In the early months of the pandemic, estimates surfaced that at least 14 million students in the United States moved to online learning (Hess, 2020).

The expansion of online courses is not merely a product of the pandemic. There has been an increasing number of higher education students enrolled in online courses in the United States since the early 2000s. A comprehensive report by the Babson Survey Research Group in 2016 indicated that more than six million students were enrolled in at least one online course. This accounted for 31.6% of all college students (Seaman, Allen, & Seaman, 2018). In 2018, more than one third (35%) of college students in the United States took at least one online course and

17% were fully enrolled online (De Brey, Snyder, Zhang, & Dillow, 2021).

The Babson study also reported that the percentage of academic leaders who rated online education as good as or better than on-ground instruction was 57.2% in 2003. This outlook of the quality of online education has shown a pattern of steady improvement from 2003 until 2012, where 77.0% of the administrators in higher education rated online as good or better. Results since then, however, have been less positive, with the results for 2015 showing only 71.4% of respondents rating online as good or better (Seaman et al., 2018).

The Noel-Levitz National Online Learners Priorities Report (2012) found that the top three priorities of the eleven enrollment factors for both undergraduate and graduate students' choosing an online course were: convenience, flexible programming, and the ability to fit education into their work schedule. The survey results were based on data from 123,594 students at 109 institutions from the fall of 2009 through the spring of 2012. This report also found that sixty-five percent (65%) of online learners perceived their experiences exceeded their expectations while twenty-four percent (24%) of them perceived their experiences met with their expectations. Seventy-three percent (73%) of the online learners were satisfied or very satisfied with their experience. In the 2017 follow-up study, the original three priorities (convenience, work schedule, and flexible pacing for completing a program) still matter and seventy-four percent (74%) of online learners were satisfied with their online programs (The Noel-Levitz National Online Learners Priorities Report, 2017).

Ortega-Maldonado, Llorens, Acosta, and Coo (2017) applied the Analysis of Variance (ANOVA) method to build student profiles of those preferring face-to-face (i.e., on-ground) versus online education. Their results indicate online Master's students were older than on-ground counterparts and were living in different cities and even countries. Unlike the on-ground students, online students did not fit a 'recent graduates' profile. Most online students had a full-time job and tended to be practitioners without too much time to spend on sustained long activities. Unlike their study's focus on Master's students in Organizations Psychology and Human Resources in one university in Spain, our research focuses on Computer Information Systems (CIS) students at both the undergraduate and graduate levels across three universities in the United States.

The research of Vidanagama (2016) involved 209 undergraduate students enrolled in computer-related degrees and focused on the role of technology. He applied the Technology Acceptance Model to determine if factors associated with online learning (e.g., perceived attitude, perceived enjoyment, and perceived usefulness) are affected by technology. The study shows that students in computing degrees are more satisfied with online learning when the technological environment (Learning Management System, software used in courses, etc.) performs adequately and is easy to use. It can be inferred from this study that students in computing fields are critical – probably more than students in other degree programs – of the technological environment involved in online learning. Our research emphasizes the understanding of CIS students to further examine their characteristics.

Cluster Analysis

Cluster analysis is a statistical process wherein data are placed into groups (i.e., clusters) based on how closely each item relates to a given set of characteristics. Classification is considered the most common use of cluster analysis; subjects are separated into groups such that each subject is more similar to other subjects within its group than to subjects outside of the group (Qualtrics, 2022). The success of clustering lies in the distinctness of the clusters that result from its application; the goal is to increase the similarity of items within a group (i.e., cluster) and to increase the difference between groups (Tan, Steinbach, Karpatne, & Kumar, 2019).

Although it is getting a renewed interest within the emerging field of data science, cluster analysis is not a new concept, as it is often used to identify groups. As Scoltock (1982) noted, cluster analysis was first developed to study the fields of biology and zoology; within these fields, clusters were used to group plants and animals and to create taxonomies for the resulting groups. Since its emergence, cluster analysis has been used in a number of other industries to distinguish attributes of a large population of subjects; over the years, it has been commonly used in a variety of areas including: biology, psychology, social sciences, medicine, etc. (Tan et al., 2019). Recently, a few researchers have used cluster analysis to study pedagogy.

Handoyo, Mukhibad, Tusyanah, and Ekaningsih (2021) utilized K-Means Clustering to group the performance of lecturers based on online pedagogical practices. This study surveyed 278 lecturers at the Universitas Negeri Saemarang

and used six variables, which included course content, teaching design, video quality, teaching service, teaching evaluation, and learning effect, to measure the performance of lecturers in online learning practices during the recent COVID-19 pandemic. The research resulted in the creation of three distinct clusters relating to lecturer performance, including low, moderate, and high performance. Based on these findings, the researchers were able to make recommendations to improve lecturers' performance in online learning which included: improving competence in operating technology-based media, adjusting the learning design to the learning conditions to make the learning process more interactive and efficient, making good and interesting learning videos, establishing intense communication with students, utilizing technology-based media when conducting learning evaluations, and motivating students to become more active (Handoyo, Mukhibad, Tusyanah, & Ekaningsih, 2021).

Koh and Chai (2014) administered a pre-course survey to teachers participating in Information and Communication Technology (ICT) lesson design professional development activities. They used the results of this survey to perform a cluster analysis to categorize teachers into groups based on their self-reported technological, pedagogical, and content knowledge (TPACK). The cluster analysis resulted in two categories of pre-service and in-service teachers, respectively. From these clusters, the researchers were able to determine that the initial TPACK differences observed in teachers lead to different effects on their perceptions of TPACK development at the end of their ICT lesson design professional development session (Koh & Chai, 2014).

Mulenga and Marbán (2020) studied the "online mathematics behaviors in the context of social media applications via online learning in mathematics activities" of high school student teachers and used cluster analysis which resulted in three clusters of students. These clusters were formed using variables relating to the extent that the student teachers use the Internet (chat, Google, wikis, etc.) when completing mathematics assignments. Although there were significant mean differences in the clustering, and the student teachers within the clusters exhibited different levels of online participation with mathematics activities, the authors concluded that these prospective mathematics teachers expressed "positive attitudes toward online learning behaviors and are likely to adopt e-learning during the coronavirus outbreak" (Mulenga & Marbán, 2020).

Mehanna (2004) examined pedagogic techniques with the goal of establishing effective e-learning practices in higher education. Effectiveness of the pedagogic techniques was determined by examining the students' outcomes on the courses that were reviewed. The results of this study revealed that seven clusters of pedagogies correlated with students' grades and an educational significance for all seven of the clusters was determined. Utilizing these pedagogies in online learning may lead to the enhancement of student learning (Mehanna, 2004).

Aggarwal and Sharma (2019) studied the performance of first year students in a Masters in Computer Application (MCA) post graduate program. The authors performed k-means clustering on the university exam data of these students and arrived at five clusters. Analyzing these clusters, the authors learned that females had better academic performance than males in the first year of the program (Aggarwal & Sharma, 2019).

Perrotta and Williamson (2018) examined the relevance of cluster analysis in categorizing and measuring online education, specifically focusing on algorithms used in learning analytics. Their focus was more on the introduction of the cluster analysis - a social science methodology - to education, rather than on the profiling of online students.

Cluster analysis, as an often-deployed methodology of studying market segmentations, is applied in this research to identify student segmentations. Since the field of online education is relatively new in comparison to the field of, say, biology, there is a lack of research in group characteristics in online education using cluster analysis. This study aims to fill this gap.

3. METHODOLOGY

An online survey regarding students' perceptions of the effectiveness of various course delivery methods was administered between the years of 2016 and 2021 at three universities: one private, one state-owned, and one state-related. The survey was IRB approved at each of the three universities and QuestionPro online survey software was used to administer the survey to students enrolled in CIS courses, regardless of major.

Two-step cluster analysis was employed to answer the two research questions regarding: whether there are specific groups of students who

shared similar characteristics with regard to their attitudes toward online education, and, if so, identifying these characteristics.

Cluster analysis, or clustering, is an unsupervised machine learning task. It involves automatically discovering natural groupings in data. Unlike supervised learning (like predictive modeling), clustering algorithms in unsupervised learning only interpret the input data and find natural groups or clusters in feature space (Wilson, 2020).

The clustering algorithm is based on a distance measure that gives the best results if all variables are independent, continuous variables that have a normal distribution, and categorical variables that have a multinomial distribution. (IBM Statistics 19).

The two-step cluster analysis is a hybrid approach which first uses a distance measure to separate groups and then a probabilistic approach (similar to latent class analysis) to choose the optimal subgroup model. (Gelbard, Goldman, & Spiegler, 2007).

Silhouette score (i.e., silhouette coefficient) is a typical measure of the success of a clustering technique. It ranges from -1 to 1. A silhouette score of 1 means that the clusters are very dense and nicely separated; whereas, a silhouette score of 0 means that clusters are overlapping. A silhouette score of less than 0 means that data belonging to clusters may be incorrect.

Cluster results are considered appropriate when the silhouette score is greater than 0.2. Though 0.2 is regarded as a fair score (Boos, Wang, Karst, Hymel, & Pediatric Brain Injury Research Network, 2021), the goal in this study was to obtain a minimum silhouette score of at least 0.3 as an indication of more robust clustering. To achieve this outcome, an iterative process of eliminating variables was deployed.

SPSS 27 was used to perform a two-step cluster analysis on the data set. Rundle-Thiele, Kubacki, Tkaczynski, and Parkinson (2015) explained that two-step cluster analysis in SPSS uses the log-likelihood measure to reveal natural groupings in a data set. It forms clusters based on both continuous and categorical data (Chiu, Fang, Chen, Wang, & Jeris, 2001; Norusis, 2008). Data transformation prior to analysis is also unnecessary.

4. RESULTS

The overall survey response at three universities in years between 2016 and 2021 was nearly 700. However, some of the respondents did not complete the entire survey; therefore, the actual number of responses to each survey question varied by question. For this two-step cluster analysis unsupervised machine learning methodology, our research focused on the student respondents that answered all 34 survey questions. Since there were already over 250 survey responses with all 34 survey questions answered, the dataset for this research contains these 251 responses. Due to the lack of time to examine each and every survey question, this research did not apply statistical methods, such as imputation, to replace missing values in order to increase the size of the data set. Regardless of this caveat, the sample size of 251 still represents a robust group for valid research.

Among the 251 students who completed the survey, 36% were female and 64% were male. The majority of these survey respondents (59%) were in the age group of 18-21. The percentage of respondents in other age groups decreased, as the ages increased, as shown in Figure 1. This basic demographic information demonstrates that the sample is representative of the population of students enrolled in CIS courses at all three universities between the years of 2016 and 2021.

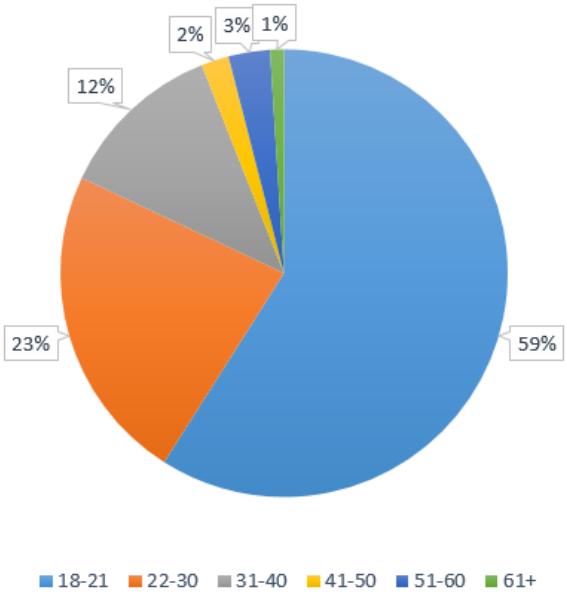


Figure 1: Percentage of Survey Respondents Grouped by Age

Research Question 1
What are the subgroup models that emerge from a multi-year, multi-university student learning preferences survey?

To answer this first research question, the following steps were taken in the data analysis.

The first step was to include all possible relevant variables in the cluster analysis. This pass included 26 relevant variables (out of 34 total variables) and did not produce any clustering, resulting in a lack of differentiation of any distinct groups or differentiated clusters.

Next, a review began to eliminate variables that were non-relevant or non-independent. This reduced the number of variables from 26 to 20. The questions where these 20 variables were extracted are shown in Appendix A. When cluster analysis was performed on these variables, two clusters were identified, as shown in Figure 2.

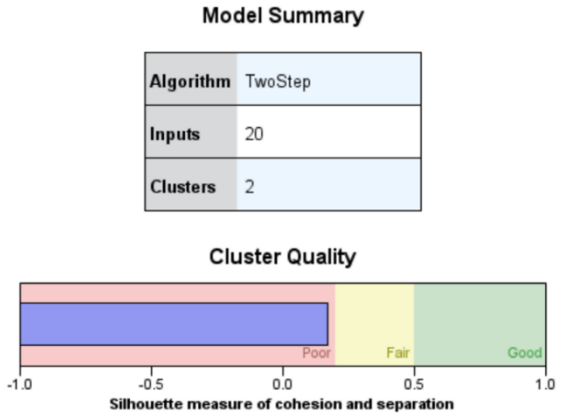


Figure 2: Clusters Obtained by Analyzing 20 Variables

The output of the 20-variable analysis depicts two clusters that were obtained from the two-step analysis. Details of the output are not illustrated in this paper due to the limitation of space. Variables were ordered based on importance, with the most important variable listed at the top. In this iteration, the most important variable was the expectation of online effort required. The first cluster had a highest selection of "less effort expected for online courses" with 68.5% expecting less effort; while the second cluster had a highest selection of "same effort expected for online and on-ground courses" with 53.2% expecting the same amount of effort. Similarly, each variable from the output can be interpreted in this fashion. The second most important variable was perception of online course effectiveness. Moving down through the list of variables, the importance of each variable becomes less in each cluster and the last two variables have no effectiveness. Since this second pass resulted in only an acceptable silhouette average of 0.2 and many variables had low or no impact, these low or no impact variables were regarded as less relevant and hence eliminated iteratively in subsequent passes in order to create more robust clusters, indicated by achieving a higher average silhouette such as 0.3.

In the next iteration, after eliminating the last two variables of the second pass that showed no effectiveness, the silhouette results remained at 0.2. Thus, variable eliminations were iteratively performed until achieving a silhouette of 0.3. This occurred when 11 variables remained as predictors.

The model summary graphic from SPSS is shown in Figure 3. The silhouette, though fair, has achieved the 0.3 goal.

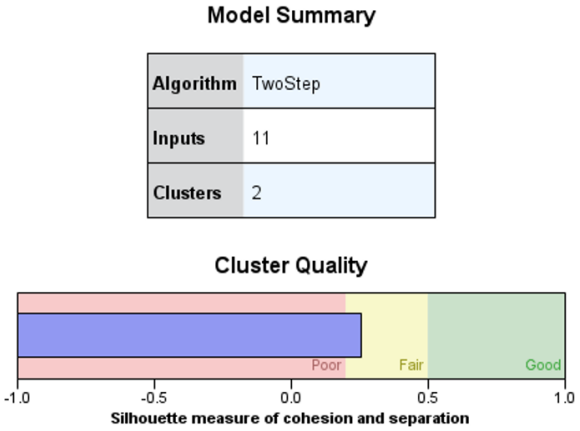
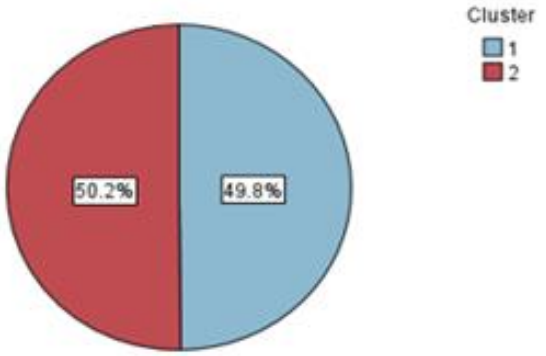


Figure 3: Clusters Obtained by Analyzing Remaining 11 Variables

As shown in Figure 4, there are two clusters of nearly equal size. These two clusters demonstrate that there are two distinct groups of students; those who prefer online education (Onliners) and those who prefer on-ground education On-grounders).



Size of Smallest Cluster	125 (49.8%)
Size of Largest Cluster	126 (50.2%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	1.01

Figure 4: Cluster Sizes

Onliners are students taking a course in which the format involves active instruction, testing, assignments, and discussion conducted over the Internet through a learning management system, in which the delivery of the course content is 100% online with no on-ground or in class meetings. Ongrounders are students taking a course on ground in the traditional brick-and-mortar classroom. Although an on-ground course might contain additional online resources such as assignments, videos, examinations, and podcasts the use of these additional resources are to enhance the class but the course is still on-ground. Finally, if the delivery format occurs when 25% - 50% of instruction, assignments, and discussions, take place online (hybrid), this online material is simply viewed as an alternative to in-person material with the intent to create a flexible learning experience.

Cluster 1 includes 49.8% of the survey respondents and represents Onliners and Cluster 2 includes 50.2% of the survey respondents and represents On-grounders.

It should be noted that during the iterative variable elimination process, some variables regarding demographics were eliminated such as gender, employment status, full-time versus part-time student status, etc. These eliminations indicated the non-significant impact of these demographics in cluster identification. The only impactful demographic variable in the remaining 11 variables was age, which will be discussed briefly later in the paper.

Research Question 2

What are the characteristics of the student subgroups that emerge from a multi-year, multi-university student learning preferences survey?

The tables displayed in Appendix B and Appendix C roughly demonstrate the clusters and the variables used to identify the characteristics in the clusters. These characteristics are listed in order of importance, with the most important variable listed first. Appendix B depicts some descriptive statistics; whereas, Appendix C provides some rough graphical presentations. Precise data analysis is provided in detail in later tables and figures.

Next, key variables were reviewed in detail, in terms of predictor importance and how they define clusters, with regard to online education. Figure 5 displays a graphical presentation of the predictor importance of each variable.

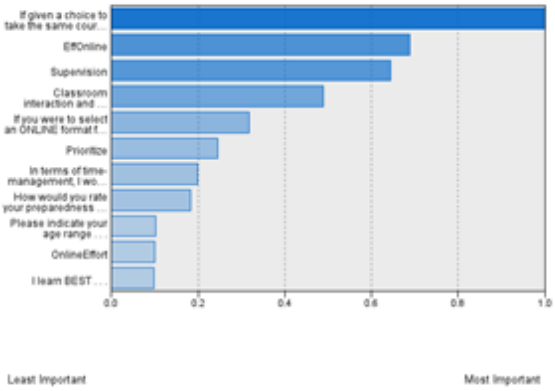


Figure 5: Predictor Importance for Each Variable

The variable with the highest predictor importance, shown in Table 1, is the question regarding selection of the online versus on-ground course format. Here, there is a clear dichotomy with 91% of the respondents in Cluster 1 preferring online and only 18% of the respondents in Cluster 2 preferring online. Hence, this reinforces the clarity of the conclusion that two discrete clusters exist in the data set, one that prefers online (i.e., Cluster 1, the Onliners)

and another that prefers on-ground (i.e., Cluster 2, the On-grounders).

If given a choice to take the same course in an ONLINE format or an ONGROUND format, would you select the ONLINE format?				
	Yes		No	
Cluster	Freq.	%	Freq.	%
1: Onliners	100	**91%**	25	18%
2: On-grounders	10	9%	116	**82%**

Table 1: Frequency and Percentage Results, by Cluster, for "Select ONLINE Format"

The second most important predictor is the rating of the effectiveness of online instruction. As shown in Table 2, Cluster 1 (the Onliners) rated online education to be effective with an average rating of 2.27; while Cluster 2's (the On-grounder's) rating of online education leaned more towards somewhat ineffective with an average rating of 3.61.

	Effective Online	
Cluster	Mean	Std.
1: Onliners	**2.27**	.928
2: On-grounders	**3.61**	1.103
Combined	2.94	1.219

Table 2: Rating Results, by Cluster, for "Effectiveness of Online Instruction"

The ability to work with or without direct supervision was the next most important predictor. As shown in Table 3, 80% of the respondents in Cluster 1 (the Onliners) indicated they work best without direct supervision compared to those in Cluster 2 (the On-grounders), of which 78% work better with direct supervision.

Work better				
	Without Supervision		With Supervision	
Cluster	Freq.	%	Freq.	%
1: Onliners	96	**80%**	29	22%
2: On-grounders	24	20%	102	**78%**

Table 3: Frequency and Percentage Results, by Cluster, for "The Ability to Work without Direct Supervision"

As shown in Table 4, classroom interaction and discussion are not essential in learning for 86% of the respondents in Cluster 1 (the Onliners) and somewhat helpful for 65% of them; whereas, classroom interaction and discussion are always helpful for 83.5% of the respondents in Cluster 2 (the On-grounders).

Classroom interaction and discussion helpfulness in learning						
	Not		Sometimes		Always	
Cluster	Freq	%	Freq	%	Freq	%
1: Onliners	24	**86%**	86	**65%**	15	16%
2: On-grounders	4	14%	46	35%	76	**84%**

Table 4: Frequency and Percentage Results, by Cluster, for "The Helpfulness of Classroom Interaction and Discussion in Learning"

The ability to prioritize also distinguished Cluster 1 from Cluster 2. As shown in Table 5, 63% of the respondents in Cluster 1 (the Onliners) can prioritize well; whereas, 73% of the respondents in Cluster 2 (the On-grounders) lack the ability to prioritize.

Prioritize				
	Can		Can't	
Cluster	Freq.	%	Freq.	%
1: Onliners	100	**63%**	25	27%
2: On-grounders	59	37%	67	**73%**

Table 5: Frequency and Percentage Results, by Cluster, for "The Ability to Prioritize"

As shown in Table 6, time management skills are much more honed for Cluster 1 (the Onliners) than Cluster 2 (the On-grounders). Fifty-seven percent of the respondents in Cluster 1 consider themselves well organized, when it comes to time management skills; whereas, over 83% of the respondents in Cluster 2 indicated that they have difficulty completing assignments and/or projects.

Time-management				
	Well-organized		Not organized	
Cluster	Freq.	%	Freq.	%
1: Onliners	117	**57%**	8	17%
2: On-grounders	88	43%	38	**83%**

Table 6: Frequency and Percentage Results, by Cluster, for "The Time Management Skills"

As shown in Appendix B and Figure 5, age is the only impactful demographic predictor for cluster identification; however, its impact was less important than the predictors discussed above in detail.

The distinctive characteristics of all 11 variables for Cluster 1 and Cluster 2 are summarized in Table 7.

Variables	**Cluster 1: Onliners**	**Cluster 2: On-grounders**
Choose online	Yes (80%)	No (92%)
Effectiveness of Online	Effective (57%)	Somewhat effective (40.5%)
Need Supervision	No (77%)	Yes (81%)
Classroom interaction importance	Sometimes helpful (69%)	Always helpful (60%)
Reason for Online	Convenience (72%)	Scheduling (56%)
Able to Prioritize	Yes (80%)	No (53%)
Time management	Well-organized (94%)	Well-organized (70%)
Preparedness for Online	Extremely prepared (29%)	Extremely prepared (6%)
Age range	18-21 (50%)	18-21 (67.5%)
Effort required for Online	Same effort (44%)	Less effort (64%)
Learn best by	Hands-on (51%)	Hands-on (74%)

Table 7: Summary of Distinctive Characteristics, by Cluster, in Descending Order of Importance, for All 11 Variables

Viewing Table 7, we find that overall, the members of Cluster 1 (the Onliners) are better organized, able to prioritize, more self-reliant, and see classroom interaction as "not essential" and only "somewhat helpful." They also tend to be slightly older, slightly less inclined to learn using hands-on methods, and believe that online courses require the same effort as on-ground courses. The Onliners view online education as effective and choose online courses for convenience; over a quarter of them feel that they are extremely prepared for online learning.

In a similar fashion, using the data in Table 7, a profile can also be built to describe Cluster 2 (the On-grounders), who view online learning as only somewhat effective. The On-grounders need supervision and consider classroom interaction important. They are less able to prioritize and less organized than the Onliners. They also tend to be younger, more inclined to learn using hands-on methods, and believe that online courses require less effort than on-ground courses. These students often choose online learning due to its ease of scheduling; however, very few of them consider themselves extremely prepared for online learning.

The above results of student segmentation regarding online education are somewhat in alignment with the employee segmentation regarding remote work. For instance, younger people have more difficulty embracing fully online education or the remote work modality.

Regarding the age demographic characteristic, the result in this research echoes previous research conducted by the authors which depicts that for those students choosing online education due to scheduling, age rather than gender, plays a significant role in choosing the online modality (Wang, Peslak, Kovacs, & Kovalchick, 2019). Deeper investigation regarding other demographics such as different age groups and generations like those applied in the Microsoft (2022) study would provide further insights.

5. CONCLUSIONS

This research begins to fill the gap in the lack of studies utilizing cluster analysis to obtain group characteristics relating to online education. Through this research the authors were able to examine student clusters and investigate distinct characteristics of students preferring online versus on-ground education.

Utilizing an iterative process of performing two-step cluster analysis of their survey data and eliminating non-relevant variables, the authors were able to arrive at two distinct student clusters in the context of online education -- Onliners and

On-grounders. The 11 variables used to create the clusters indicate the characteristics of students within each cluster.

The 11 characteristics, found in this study, can be used to build a profile of a typical online student and that of a typical on-ground student. These profiles can be used by decision makers in higher education when making policies and allocating resources. For instance, this research suggests post-graduate programs embrace more online education than undergraduate programs.

These profiles can also be used in strategic planning with regard to how to market, recruit, and retain students for both online and on-ground educational programs. For instance, online post-graduate programs can be better marketed to employees who have already adopted fully-remote work.

The identification of specific online versus on-ground clusters and their identifying characteristics provides important insights to better understand students and also to better assist them in improving their acceptance and performance of online education, when necessary. Who knows what the future will bring – another pandemic or a climate change disaster could move education 100% online again. It is better to be prepared for the unknown.

This survey was limited in its audience to only those students enrolled in a CIS course at one of three universities. Therefore, one may conjecture that the majority of survey respondents were computing majors or students with some computing background. Surveying students enrolled in a variety of general education courses (humanities, fine arts, social sciences, etc.) would likely result in a more well-rounded characterization of students who prefer online versus on-ground education.

Now that profiles have been built to describe students enrolled in CIS courses who prefer online versus on-ground education, the authors can continue this study by making changes to the marketing strategies within their departments based on these profiles. Future research may examine the effects of these marketing changes by surveying students recently enrolled in these online and on-ground programs.

6. REFERENCES

Aggarwal, D. & Sharma, D. (2019). Application of Clustering for Student Result Analysis. *International Journal of Recent Technology and Engineering (IJRTE)*, 7(6C), 50-53.

Bishop, K. (2022). Is remote work worse for wellbeing than people think? *BBC*. June 17, 2022. Retrieved August 29, 2022 from https://www.bbc.com/worklife/article/20220616-is-remote-work-worse-for-wellbeing-than-people-think

Boos, S. C., Wang, M., Karst, W. A., Hymel, K. P., & Pediatric Brain Injury Research Network (PediBIRN) Investigators. (2022). Traumatic head injury and the diagnosis of abuse: a cluster analysis. *Pediatrics*, 149(1).

Chiu, T., Fang, D.-P., Chen, J., Wang, Y., & Jeris, C. (2001). A robust and scalable clustering algorithm for mixed type attributes in large database environment. *Proceedings of the 7th ACM SIGKDDD International Conference on Knowledge Discovery and Data Mining*, ACM SIGKDDD, San Francisco, CA, 263-268.

De Brey C., Snyder T. D., Zhang A., Dillow S. A. (2021). *Digest of Education Statistics 2019* (NCES 2021-009). National Center for Education Statistics.

Gelbard, R., Goldman, O., & Spiegler, I. (2007). Investigating diversity of clustering methods: An empirical comparison. Data & Knowledge Engineering, 63(1), 155-166.

Handoyo, E., Mukhibad, H., Tusyanah, T. & Ekaningsih, L. (2021). Lecturers' Performance in Pandemic Era Based on Online Pedagogical Practices in Universitas Negeri Semarang (UNNES), Indonesia: A Cluster Analysis-Based Approach. *Education Sciences: Theory & Practice (JESTP)*, 21(4), 138-154.

Hess A. (2020, March 26). *How coronavirus dramatically changed college for over 14 million students*. CNBC.com. https://www.cnbc.com/2020/03/26/how-coronavirus-changed-college-for-over-14-million-students.html

IBM SPSS Statistics 19 Statistical Procedures Companion

Koh, J. H. L., & Chai, C. S. (2014). Teacher clusters and their perceptions of technological pedagogical content knowledge (TPACK) development through ICT lesson design. *Computers & Education*, 70, 222-232.

Lufkin, B. (2022). The companies doubling down remote work. *BBC.* July 24, 2022. Retrieved August 29, 2022 from https://www.bbc.com/worklife/article/20220

722-the-companies-doubling-down-on-remote-work

Markman, A. (2021). Why you may actually want to go back to the office. *Harvard Business Review*. July 1, 2021. Retrieved August 29, 2022 from https://hbr.org/2021/07/why-you-may-actually-want-to-go-back-to-the-office

Mehanna, W. N. (2004). e-Pedagogy: the pedagogies of e-learning. *Research in Learning Technology*, 12(3), 279-293.

Microsoft. (2022). Great expectations: making hybrid work. Retrieved August 29, 2022 from https://www.microsoft.com/en-us/worklab/work-trend-index/great-expectations-making-hybrid-work-work

Mulenga, E. M. & Marbán, J. M. (2020). Prospective Teachers' Online Learning Mathematics Activities in the Age of COVID-19: A Cluster Analysis Approach. *EURASIA Journal of Mathematics, Science and Technology Education*, 16(9).

Noel-Levitz, (2012). Noel-Levitz national online learners priorities report. Retrieved from https://www.ruffalonl.com/upload/Papers_and_Research/2012/2012_Online_Leaners_Report.pdf.

Noel-Levitz, (2017). Noel-Levitz national student satisfaction and priority report. Retrieved from https://learn.ruffalonl.com/rs/395-EOG-977/images/2017_National_Student_Satisfaction_Report_1.0.pdf

Norusis, M.J. (2008), SPSS 16.0 Guide to Data Analysis, 2nd ed., Prentice Hall.

Ortega-Maldonado, A., Llorens, S., Acosta, H. & Coo, C. (2017). Face-to-face vs on-line: an analysis of profile, learning, performance and satisfaction among post graduate students. *Universal Journal of Educational Research*, 5(10), 1701-1706.

Peslak, A., Kovalchick, L., Wang, W. & Kovacs, P. (2021). Effectiveness of educational delivery modes: a study in computer information systems. *Journal of Information Systems Education*, 32(4), 253-261.

Perrotta, C. & Williamson, B. (2018). The social life of learning analytics: cluster analysis and the 'performance' of algorithmic education. *Learning Media and Technology*, 43(1), 3-16.

Qualtrics. (2022). What is cluster analysis? When should you use it for your survey results? (2022). Retrieved August 29, 2022 from https://www.qualtrics.com/experience-management/research/cluster-analysis/

Rundle-Thiele, S., Kubacki, K., Tkaczynski, A., & Parkinson, J. (2015). Using two-step cluster analysis to identify homogeneous physical activity groups. *Marketing Intelligence & Planning* 33, 522-537.

Scoltock, J. (1982). A survey of the literature of cluster analysis. *The Computer Journal*, 25(1), 130-134.

Seaman, J. E., Allen, I., E., & Seaman, J. (2018). *Grade Increase: Tracking Distance Education in the United States*. (ED580852). ERIC. https://files.eric.ed.gov/fulltext/ED580852.pdf

Shevlin, R. (2022). JPMorgan Chase gets "work from home" all wrong. *Forbes*. August 22, 2022. Retrieved August 29, 2022 from https://www.forbes.com/sites/ronshevlin/2022/08/22/jpmorgan-chase-gets-work-from-home-all-wrong/?sh=1badf24a7cb0

Smith, M. (2022). Twitter, Reddit and 8 other companies offering permanent remote or hybrid work—and hiring right now. *CNBC*. April 13, 2022. Retrieved August 29, 2022 from https://www.cnbc.com/2022/04/13/10-companies-that-switched-to-permanent-hybrid-or-remote-work-and-hiring-right-now.html

Stewart, J., Miller, M., Audo, C., & Stewart, G. (2012). Using cluster analysis to identify patterns in students' responses to contextually different conceptual problems. *Physical Review Special Topics-Physics Education Research*, 8(2), 020112.

Tan, P.N., Steinbach, M., Karpatne, A. & Kumar, V. (2019). Cluster Analysis: Basic Concepts and Algorithms. *Introduction to data mining* (2, pp. 525-612). New York, NY: Pearson. Retrieved June 15, 2022 from https://www-users.cse.umn.edu/~kumar001/dmbook/ch7_clustering.pdf

Vidanagama, D. U. (2016). Acceptance of E-learning among undergraduates of computing degrees in Sri Lanka. *International Journal of Modern Education and Computer Science,* 8(4), 25-32.

Wang, W., Peslak, A., Kovacs, P., & Kovalchick, L. (2019). What really matters in online education, *Issues of Information Systems*, 20(1), 40-48.

Wilson, J. (2020). How are clustering algorithms different from supervised learning? *Technical-QA.COM*. Retrieved June 15, 2022 from https://it-qa.com/how-are-clustering-algorithms-different-from-supervised-learning/

APPENDIX A

Survey Questions Used for 20 Variable Cluster Analysis

Note: the number of the question refers to the number in the original survey which contains 34 variables

2) If given a choice to take the same course in an ONLINE format or an ONGROUND format, would you select the ONLINE format?

☐ Yes
☐ No

3) If you did select an ONLINE format for a course, what would be the main reason?

☐ Convenience
☐ Scheduling
☐ Delivery Method
☐ To take a particular professor
☐ Other (please specify)
If you selected other, please specify_____

4) I have taken (or am currently taking) a course that is completely online or is partially online.

☐ Yes
☐ No

6) What type of formal training did you receive to prepare you to take an online course?

☐ No formal training received
☐ Training and documentation provided by my school
☐ Self-trained
☐ Training from course instructor or other faculty member
☐ Training from another student
☐ Other (please specify)
If you selected other, please specify_____

7) How would you rate your preparedness (to take an online course) prior to taking your online course?

☐ Extremely unprepared
☐ Somewhat unprepared
☐ Neither unprepared nor prepared
☐ Somewhat prepared
☐ Extremely prepared

8) Do you perceive the OVERALL effectiveness of courses that are offered COMPLETELY online as…

☐ Very effective
☐ Effective
☐ Somewhat effective
☐ Somewhat ineffective
☐ Ineffective
☐ Very ineffective

9) Do you perceive the OVERALL effectiveness of courses that are offered PARTIALLY online and

PARTIALLY onground (i.e., Hybrid) as…

☐ Very effective
☐ Effective
☐ Somewhat effective
☐ Somewhat ineffective
☐ Ineffective
☐ Very ineffective

10) Do you perceive the OVERALL effectiveness of courses that are offered ONGROUND but have an ONLINE SUPPLEMENT (i.e., online materials provided on BlackBoard or on an instructor's website) as…

☐ Very effective
☐ Effective
☐ Somewhat effective
☐ Somewhat ineffective
☐ Ineffective
☐ Very ineffective

15) Select one of the following choices

☐ I work better without direct supervision
☐ I work better when someone is there to keep me focused

16) Select one of the following choices

☐ I can prioritize my own workload
☐ I tend to put work off until later

17) Select one of the following choices

☐ I would allocate as much time and effort for an online course as I would for an on-ground course
☐ I feel that LESS time and effort is required for an online course (as compared to an on-ground course)
☐ I feel that MORE time and effort is required for an online course (as compared to an on-ground course)

18) In terms of time-management, I would describe myself as…

☐ Well organized
☐ Having difficulty completing assignments and/or projects

19) Classroom interaction and discussion is…

☐ Not essential for me to learn/understand
☐ Sometimes helpful for me to learn/understand
☐ Always helpful for me to learn/understand

20) Which of the following aspects could influence my decision to take an online course…

☐ Instructor teaching the course
☐ Design of the course
☐ Subject matter of the course
☐ Other (please specify)
If you selected other, please specify_____

22) I learn BEST…

☐ By seeing (visually)
☐ By listening
☐ By reading
☐ By doing (hands-on)

26) Are you enrolled as a …

☐ Full-time student
☐ Part-time student

27) Which of the following best describes your living arrangement…

☐ Resident student (live on campus)
☐ Commuter student (live off campus)

32) Are you currently employed as a…

☐ Full-time employee (>40 hours/week)
☐ Part-time employee (
☐ Not currently employed

33) Please indicate your sex…

☐ Male
☐ Female

34) Please indicate your age range…

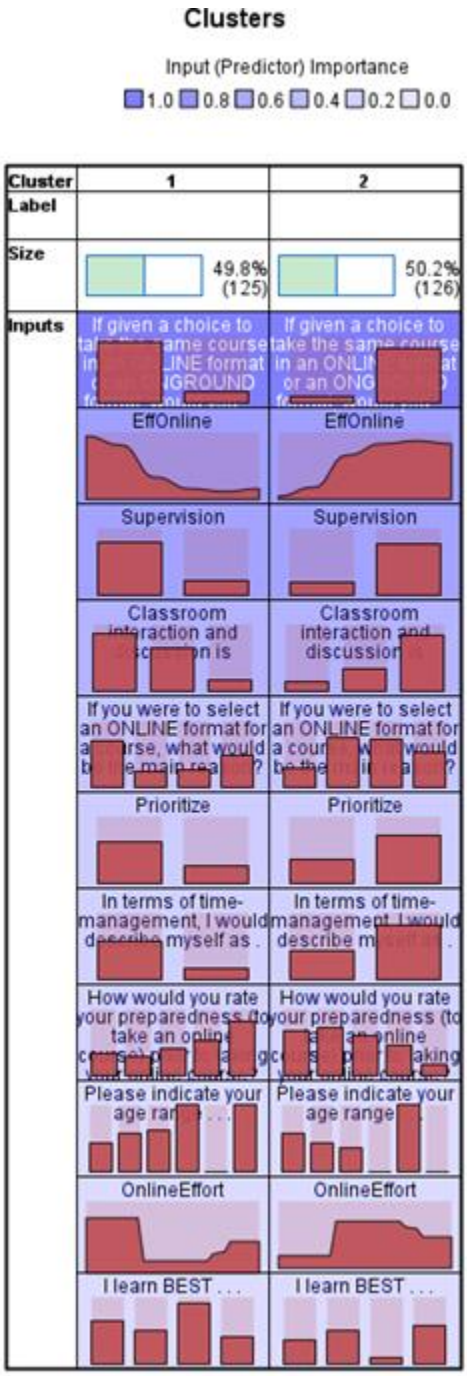
☐ 18 - 21
☐ 22 - 30
☐ 31 - 40
☐ 41 - 50
☐ 51 - 60
☐ 61 or older

APPENDIX B

Variables Identifying the Characteristics of Each Cluster, Listed in Order of Importance with Some Descriptive Statistics

Cluster	1	2
Label		
Size	49.8% (125)	50.2% (126)
Inputs	If given a choice to take the same course in an ONLINE format or an ...	If given a choice to take the same course in an ONLINE format or an ...
	EffOnline 2.27	EffOnline 3.61
	Supervision 1 (76.8%)	Supervision 2 (81.0%)
	Classroom interaction and discussion is	Classroom interaction and discussion is
	If you were to select an ONLINE format for a course, what would be the main reason?	If you were to select an ONLINE format for a course, what would be the main reason?
	Prioritize 1 (80.0%)	Prioritize 2 (53.2%)
	In terms of time-management, I would describe myself as .	In terms of time-management, I would describe myself as .
	How would you rate your preparedness (to take an online course) prior to ...	How would you rate your preparedness (to take an online course) prior to ...
	Please indicate your age range . . .	Please indicate your age range . . .
	OnlineEffort 1.85	OnlineEffort 2.15
	I learn BEST . . . 4 (52.8%)	I learn BEST . . . 4 (73.8%)

APPENDIX C
Relative Frequencies of Responses to Each Question/Characteristic in Graphical Form



Teaching Case

Design Thinking: Facilitating Consumer Access to Community Services

Jason Ferguson
jasferguson@bellevue.edu
College of Science and Technology
Bellevue University
Bellevue, NE

Michelle Louch
louchm@duq.edu
John G. Rangos, Sr., School of Health Sciences
Duquesne University
Pittsburgh, PA

Abstract

The case focuses upon 1) a public health outcome, namely improving access to healthcare, 2) systems and design thinking approaches to software development and the internet of things, 3) mockup tools and user interface design, 4) understanding stakeholder requirements and feature requests, 5) presentation of a prototype application. This case presents a blend of healthcare management and technology concerns and is an appropriate capstone project for an undergraduate information systems course. Depending on the instructor's individual academic needs, the various assignments can also be modified for lower-level courses. The tasks were originally designed by one of the authors for a course called Creating Solutions with Integrated Technology and are used with permission.

Keywords: Design Thinking, Healthcare Outcomes, Internet of Things, Population Health Management, Public Health, Systems Thinking

1. THE PROBLEM

As part of a public health initiative, this August, McKeesport Mercy Hospital (MMH) is offering its second annual Community Health Clinic which will provide free physicals and vaccines for all ages. Last year's event was relatively successful, and the hospital's board is hoping to expand and build on it this year.

Greg is the head of marketing at MMH and has been charged with getting the word out so that as many residents as possible are able to take advantage of the clinics. He and his team have developed several eye-catching advertisements for bus stops, billboards, social media, television, and flyers. They even added an option on the MMH website that allowed people to schedule their appointments electronically.

At yesterday's meeting, just before they wrapped up, Greg asked if anyone had any questions. Cara, the newest member of the team, spoke up. "I was looking at the feedback on last year's clinics, and a number of comments related to access and transportation," she said. "I know that this year we changed the hours to better accommodate working adults, which is going to

be a great help, but we aren't doing anything about transportation."

Before Greg could reply, other team members chimed in, dismissing Cara's concern.

"We can't do anything about transportation. It's up to them to find their way here."

"Exactly. There's public bussing, Uber, Lyft, and even jitneys."

"We're advertising at the bus stops. And we do have a shuttle, you know."

By the third comment, Greg held up his hand. "Let Cara finish." He looked at her. "Go ahead, what are you thinking?"

"The MMH shuttle is a good start in terms of providing a way out and back, but it's pretty limited since it only runs in a loop from the strip mall to the hospital. We know that there are other transportation options out there, but it's possible that not everyone knows what is or is not available. Or they may know about them, but they aren't sure how to hire an Uber or how to find a bus schedule or get to the mall.

"We added an app to the website to make it easy for people to schedule their appointments, we have QR codes on all of the print media, and we have a dedicated phone line for those people who don't want to use tech to make their appointments."

"What if we add an app that helps people find transportation? Like 'click here if you need a bus schedule' or 'click here is you want our shuttle schedule' or 'click here to go to the Uber website and set up an account.' There's a lot of possibilities here.

"Part of MMH's mission statement relates to outreach and helping the underserved. If we're going to do that, we need to recognize that one barrier for some people getting the healthcare that they need is reliable transportation."

Finished, Cara sat back and let what she said sink in. After a moment, Greg nodded. "You're right. An app that makes it easier for people to plan their trip can't hurt. We need to get our numbers up. The hospital is looking at public health outcomes and income, so the more patients we have, the better."

He looked at the team. "Go back over the demographics, look at our service area, and come up with some ideas on how we can make it easier for people to get here. We'll meet Tuesday after lunch to figure out what we need our software development team to do."

2. DEFINING THE APPLICATION

Tuesday's meeting went without a hitch, and Greg was able to give software development team a clear picture of what they wanted in the application.

Application Requirements:

- Must be cross-platform
- Must be user friendly – the fewer the clicks, the better
- Should be intuitive but also easy for those anxious about tech ~~to~~ use
- Mapping option so users can see how to get to MMH from their homes – link to Google's map application but autofill starting and ending destinations?
- Links to outside transportation apps such as Yellow Cab, Uber, Lyft, and other local options
- Link to MMH shuttle – offer pick-up and drop-off times as well as locations (MMH & mall)
- Link to public transportation system
    - Specifically, link to busses that pick-up and drop-off at the mall and the hospital
    - Make sure that users know what bus fares will be, if they can pay in cash, or if a "electronic ticket" is required/an option

Features:

- Provide "best route" recommendations for users
- Recommend best travel options based on time of day
- Provides users a way to provide feedback on their experience with the application in finding transportation to a MMH health clinic

3. FIRST LAB ASSIGNMENT

View the following videos. The first provides a short overview of design thinking. The second and third showcase examples of successes and failures in both systems design and design thinking approaches to solving real-world problems:

- AJ&Smart's explanation: *What is Design*

Thinking? An Overview. (https://youtu.be/gHGN6hs2gZY)
- Doug Dietz's TEDx talk: *Transforming healthcare for children and their families* (https://youtu.be/jajduxPD6H4)
- Elisabeth McClure's TEDx talk: *Are children really more creative than adults?* (https://youtu.be/g00o6LCmaMI)

Next, visit the landing page of the design-thinking website that provides a short explanation as to what design thinking is: https://en.dt-toolbook.com/.

Discuss why, or why not, some ideas work and others fail.

At the end of the lesson, you will be divided into teams of 3 or 4. Review the provided case once more, and produce a draft of what your team believes are:

1. The boundaries of the system/solution they will be creating a prototyping.
2. The operating environments of the system/solution.
3. The system components and interactions the system/solution will support.

4. SECOND LAB ASSIGNMENT

Next, you team will plan to conduct interviews to understand and examine the problem scenario from the end-user point of view. Your instructor will provide information regarding these interviews.

The overall objective of this assignment is to help your team develop an in-depth understanding of the problem.

To facilitate the organization and planning process, you are encouraged to use the Ask 5x Why questions found on *The Design Thinking Playbook* website (https://en.dt-toolbook.com/tools). The provided link takes you to *DT Tools.* Once there, you can download the "basic PDF template" *05_Question 5X Why* for free.

5. THIRD LAB ASSIGNMENT

This assignment asks teams to continue exploring the factors/items required for successful completion of the MMH project.

This assignment provides a more structured approach by asking the students to develop a more specific interview plan based on the 5W + H model.

The 5W + H model focuses on identifying the following: what, who, why, where, when, and how. Consider using the free PDF template, *06_5WH Questions*, provided on *The Design Thinking Playbook* website (https://en.dt-toolbook.com/tools).

- **What**: What issue or problem are you exploring? What will your final result look like?
- **Who**: Who is (are) the user(s) you have identified?
- **Why**: Why is this an issue? Why it is important to solve?
- **Where**: What is the 'environment' this will be used. Ex. Web page, mobile application, kiosk, etc.
- **When**: When will the target market use this? What is our timeline?
- **How**: How will you solve the issue? Be sure to include at least four (4) actionable steps that you intend to take.

Your team will complete a 5W+H template and fully address each of the following for the issue or problem presented: what, who, why, where, when, and how.

Once the worksheet is completed, another interview is recommended.

6. FOURTH LAB ASSIGNMENT

This lab is designed to help your team make decisions, remove "analysis paralysis" barriers, resolve any disagreements, and focus on the best ideas toward a design solution for the project.

Your team will develop a minimum of five (5) "How Might We?" statements.

It is recommended that team members pay particular attention to the difference between requirements and features while brainstorming to ensure that the team is aligning solutions to the problem(s) identified.

Use the Planning Matrix provided in Appendix A. Your team should maintain and update the planning schedule throughout the remainder of the project.

7. FIFTH LAB ASSIGNMENT

At this point, review the project requirements, the problem identified, and the priorities that your team identified in the previous lab and complete the following matrix with components and requirements for the project (Appendix B).

Your team should maintain and update this matrix as necessary.

Once the matrix is complete, create a prototype of the application, following these steps:

1. Begin creating the physical prototype for the project.
2. Review and update their planning schedule and submit a completed version by the end of the week.
3. Meet with the instructor or problem sponsor to receive feedback and to answer questions.
4. Continue to make incremental improvements to your project.
5. Review and update the planning schedule.
6. Repeat steps 3-5 above for (3-week limit) until the final prototype solution is ready to be presented.

8. FINAL LAB ASSIGNMENT

Each team will create a presentation to highlight their project solution. They should target their presentation to the marketing team from MMH.

The following structure is recommended:

1. For the introduction, include the following:
    a. Team Members + Photos
    b. Degree/Major
    c. Links to LinkedIn Profiles
    d. How each team member's experience and/or background is relevant to the problem
    e. Pictures of the problem sponsor.
2. Overview of your process: Problem Identification, Interviews, Observations.
3. Why is it important to solve this problem? – Include any research findings.
4. Planning Process.
5. Prototype Iterations – Include pictures of your progress.
6. Final Iteration (Where did you end up? What gaps remain?) – Pictures/Videos of the final iteration.
7. Lessons learned.
8. Demo of the final iteration.
9. Questions from the audience.

Presentations should be no longer than 30 minutes.

9. DISCUSSION QUESTIONS

In addition to questions covered during the final presentations, your instructor may ask some of the questions below.

Healthcare- and business-related questions:

- Imagine that you want to monetize this application. What businesses, other than healthcare, might find this useful and why? (SA)
- What are the benefits/drawbacks to offering applications for little to no charge? Why? (SA)
- If you choose to monetize this application, would you charge the end user or use ads to keep it free? Why? (SA)
- Summarize your key takeaways from this case study and describe how you have applied or plan to apply them in the business or healthcare field. (SA)
- Applications have little value if the targeted users are unaware of them. How would you market this particular application? Use Canva to create an advertisement that for your application (See Resources for links to Canva and free images). Be sure to include a QR code that can take them directly to the application itself.

IT- and IS-related questions:

- Are there other problems that an application such as this can solve?
- What are the advantages and disadvantages of a design-thinking approach and a systems-thinking approach? (SA)
- Imagine that you have unlimited money and time, what could you do to make this application better?
- Summarize your key takeaways from this case study and describe how you have applied or plan to apply them in the technical field. (SA)

Questions marked with an *SA* can be used as short-answer essay questions. Quality responses will demonstrate critical thinking and an

understanding of the skills required to create the application and meet the project sponsor's requirements.

10. RESOURCES

Prototype/Mockup Tools recommended for this assignment:

- Balsamiq: https://balsamiq.com
- Invision: https://www.invisionapp.com/home
- Mockplus: https://www.mockplus.com
- Prototypr.io: https://prototypr.io/toolbox

Links for the First Lab Assignment:

AJ&Smart. (2020). *What is Design Thinking? An Overview.* Retrieved on March 17, 2023. https://youtu.be/gHGN6hs2gZY

dTP. (2021). *Intro to Design Thinking.* Retrieved January 10, 2023. https://en.dt-toolbook.com/

TEDx. (2012, May 19). *Transforming healthcare for children and their families: Doug Dietz at TEDxSanJoseCA* [Video]. YouTube. https://youtu.be/jajduxPD6H4

TEDx. (2019, January 8). *Are children really more creative than adults? Elisabeth McClure TEDxAarhus* [Video]. YouTube. https://youtu.be/g00o6LCmaMI

Link for the Second Lab Assignment:

dTP. (2021). *DT Tools: 05_Question 5x Why.* Retrieved June 13, 2022, from https://en.dt-toolbook.com/tools

Link for the Third Lab Assignment:

dTP. (2021). *DT Tools: 06_5WH Questions.* Retrieved June 13, 2022, from https://en.dt-toolbook.com/tools

Links for Canva assignment:

- Canva: http://www.canva.com
- The Noun Project (free images): https://thenounproject.com/
- Pixabay (free images): https://pixabay.com

APPENDIX A
Planning Mode

Feature/Requirement	Task Name	Description and Details	Who's Responsible	Status (OK, Doing, To Do)	Remaining Effort

APPENDIX B
Requirements Matrix

Project Requirements	Components Needed