EasyChair Preprint

№ 10922

Exploring the Effectiveness of Coding Education Methods and Their Impact on the Data Science Job Market Requirements – a Systematic Review

Nipuna Thalpage and Rohit Bansal

September 19, 2023

Chapter 8

Exploring the Effectiveness of Coding Education Methods and their Impact on the Data Science Job Market Requirements – A Systematic Review

Nipuna Sankalpa Thalpage[1], Rohit Bansal[*2], PhD,

[1]Cardiff School of Technologies, Cardiff Metropolitan University, Cardiff, Wales
[2]Department of Management Studies, Vaish College of Engineering, Rohtak, India

Abstract

The growing demand for data science and coding skills has led to an increase in the number of educational programs in these areas. However, the effectiveness of these programs in terms of career development outcomes remains unclear. This systematic review aims to evaluate the available literature on the impact of coding education on data science career development. The study will follow a rigorous and systematic approach to identify and analyze relevant research on data science and coding education. A comprehensive search of relevant databases will be conducted using keywords and MeSH terms related to data science, coding education, career development, and program evaluation. The purpose of this systematic review is to evaluate the available literature on the effectiveness of coding education for Data science career development. The increasing demand for professionals with data science and coding skills has led to a growth in the number of educational programs in these areas. However, the quality and effectiveness of these programs are not well understood. This systematic review aims to address this gap by synthesizing the existing research on the impact of coding education on data science career development.

Keywords: systematic review, coding education, Data Science, Career Development

Introduction

Data science has become increasingly important in various industries due to the growing demand for professionals with these skills. The field of data science, in particular, has been identified as a key driver of innovation and growth in many sectors (Davenport and Patil, 2012). As data becomes more prevalent in business and society, the ability to analyze and interpret it is essential for success in many career paths (Manyika et al., 2011).

Corresponding Author's Email: rohitbansal.mba@gmail.com.

The demand for individuals with coding skills has led to a proliferation of educational programs in these areas. In recent years, universities and other institutions have created a variety of programs aimed at teaching these skills, ranging from coding boot camps to full-time degree programs as shown in Figure 1. However, the quality and effectiveness of these programs are not well understood, and there is a need for research to evaluate their impact on career development.

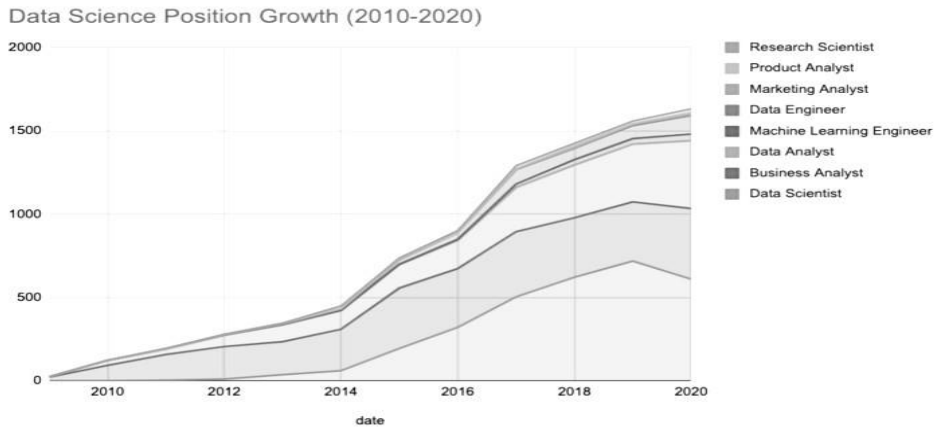

Figure 1. Expansion of Data Science Careers (Glen, 2021).

Importance of Coding Education for Data Science

Coding education is increasingly important for individuals seeking to develop their skills and advance their careers. The job market for professionals with these skills is growing rapidly, with employers across a variety of industries seeking candidates with expertise in data science. Furthermore, individuals with data science and coding skills tend to have higher salaries and greater job security than those without these skills (Karaoglu et al., 2022).

In addition to the potential benefits for individuals, data science skills are also important for the overall success of organizations. The ability to analyze and interpret data can provide insights that drive innovation and growth, leading to a competitive advantage for businesses (Manyika et al., 2011). Therefore, investing in coding education for Data Science can be beneficial for both individuals and organizations.

Purpose of the systematic review

The purpose of this systematic review is to evaluate the available literature on the effectiveness of coding education for data science career development. Specifically, the review aims to answer the following research questions.

Research questions

1: What are the factors and methods that contribute to the effectiveness of coding education programs?

2: What are the main skills and requirements of the Data Science job market?

Literature Review

The body of research on coding education for data science career development is extensive and covers a wide range of disciplines. In a study (Miller, 2019) study focused on the promotion of mathematical knowledge and thinking, with a specific focus on the identification of mathematical patterns and structures, through engagement with coding activities. Similarly, (Feijóo-García et al., 2020) evaluated the performance of adult participants with no prior coding experience on conventional block-based programming platforms. Several studies have contributed to our understanding of key aspects of the field of coding education and related areas. (Ari et al., 2022) conducted research on early childhood preservice teachers' perceptions of computer science and gender stereotypes. Their findings shed light on the importance of addressing gender biases and stereotypes to promote inclusive computer science education. In line with this, the diverse range of research on data science and coding education highlights the significance of this topic in modern-day education and the need for further investigation into its effectiveness in career development.

Furthermore, (Fischer et al., 2021) explored the impact of coding platforms such as Codewars and CodinGame on undergraduate students' motivation in coding education. Their study revealed a positive effect on student motivation, emphasizing the potential of such platforms to enhance learning experiences. Additionally, (Hundhausen et al., 2013) presented a set of best practices for implementing Pedagogical Code Reviews (PCR) in software engineering education. Their work provides valuable insights for educators teaching PCRs, offering guidance on effective teaching practices in this area.

(Prato, 2017) explores the use of programming logic and coding languages as educational tools for young children, elementary and middle school children, and adults. The study emphasizes the importance of providing access to coding education at an early age to foster technological literacy and promote career opportunities in the tech industry. (Duda et al., 2021) aimed to cover basic coding topics and data science concepts through live coding and hands-on practice. Furthermore, (Emery et al., 2021) provide an important window into how data science is taught in higher education biology programs and how instructors across disciplines can best move forward to empower their students.

Another study conducted by (Karahan et al., 2015) suggested that media design processes could serve as an effective instructional approach in STEM education. The research findings indicated positive impacts on students' attitudes toward science and technology classes.

The growing importance of data science has captured considerable interest in various industries, leading organizations to prioritize the establishment of their data science capabilities. Nonetheless, practitioners often face obstacles when it comes to creating technology-driven activities based on data (Kumari et al., 2022). In a related context, (Pournaras, 2017) sheds light on the insights gained from designing and instructing a unique

data science course that encompasses multiple disciplines. The study emphasizes the significance of collaboration among stakeholders and the implementation of a comprehensive curriculum to address these challenges effectively. By drawing attention to the practical experiences shared by (Pournaras, 2017) and the difficulties identified by (Kumari et al., 2022), a clearer connection emerges, underscoring the importance of addressing the barriers in technology-driven data science activities through collaborative efforts and robust educational frameworks.

Mikroyannidis et al., (2018) propose a system for enhancing data science accreditation using blockchain technology. In contrast, (Oesterreich et al., 2019) argue that there is no evidence to suggest that data science skills are essential. Instead, they study the impact of digitization on the controller's job roles, skills, and competencies.

Mikalef et al., (2019) propose an explanatory model to understand the gap between intentions and actions regarding data science. They build their model on survey responses from key executives in Norwegian firms. (Horton and Hardin, 2021) discuss ways to teach computational thinking in the special issue of the Journal of Statistics and Data Science Education.

This literature review provides evidence for the effectiveness of data science and coding education in preparing individuals for career development. The literature suggests that traditional data science programs and online platforms such as Codecademy and Udemy have been effective in increasing coding skills and career opportunities. Additionally, the study of Dart and Flutter has gained popularity and is effective in increasing job opportunities and salaries. However, the review also highlights the challenges and barriers to the successful implementation of coding education programs for the data science job Market. These challenges include a lack of access to resources, difficulty in finding qualified instructors, and a rapidly changing industry landscape that requires continuous adaptation of curriculum and teaching methods. To address these challenges, best practices have been identified for the effective implementation of data science and coding education programs, including incorporating practical projects and real-world applications, promoting collaboration and peer learning, and continuous evaluation and improvement of the curriculum.

Further, considering the pieces of literature on coding education for data science career development, there is still a gap in exploring the effectiveness of coding education methods and their impact on the data science job market requirements. While the literature has covered a wide range of disciplines and topics, including laboratory tasks, argumentation, information systems curricula, and interdisciplinary approaches, there is a need for further investigation into best practices for implementing coding education in data science. Several studies have proposed different methods, such as live coding and hands-on practice, media design processes, web scraping, and natural language processing, but there is a lack of research on the outcomes and impact of these approaches on the data science job market requirements. Understanding the effectiveness of coding education methods in preparing individuals for the data science job market is crucial for educators, employers, and job seekers alike. Finally, these kinds of literature also underscore the importance of coding education in preparing individuals for data science and careers in the current job market. As technology continues to advance and data-driven decision-making becomes increasingly important across industries, educational

institutions and training programs must develop and implement effective programs that meet the needs of the industry.

Methodology

To conduct this systematic review, we followed a streamlined approach that involved four stages:

1. Identifying relevant studies
2. Screening the studies based on eligibility criteria
3. Extracting data from the selected studies
4. Synthesizing the results.

First, the author conducted a comprehensive search of two databases: Google Scholar and IEEE Explore which are relevant to the field of data science education. We used a combination of keywords and MeSH terms such as "data science education", "coding education", "programming education", "data literacy", "data skills", "curriculum development", "industry needs", and "job market". The search was limited to articles published in English between 2010 and 2022 that focused on data science and coding education.
Secondly, we screened the articles based on our inclusion and exclusion criteria.

The inclusion criteria were
1. Studies focused on data science and/or coding education
2. Studies examined the effectiveness of educational programs or curriculum development
3. Studies reported on the skills and knowledge required for data science
4. Studies that are published in a peer-reviewed journal.

The exclusion criteria were
1. Studies that focused on other subject areas than data science or coding education.
2. Studies that were not focused on education, curriculum development, or data science,
3. Studies that were not peer-reviewed,
4. Studies that were published before 2010.

Third, we used thematic analysis, which is a technique used to recognize, examine, and interpret patterns or themes that convey significance within qualitative data (Braun and Clarke, 2016) to analyze the selected studies. We organized the data into themes and patterns based on the research questions and objectives. Themes were developed through a process of coding and categorization of the data, with a focus on identifying patterns and relationships in the data. We used both inductive and deductive approaches to ensure that the themes were grounded in the data but also aligned with our research objectives.

Lastly, we synthesized the results of the selected studies and presented them in a narrative synthesis. We identified common themes and patterns across the studies and summarized them. By using thematic analysis, we aimed to provide a comprehensive and meaningful

interpretation of the data that can inform future research and practice in the field of data science and coding education.

Although we have used PRISMA guidelines, A method that involves a well-defined and unambiguous research question and utilizes explicit and systematic methods to identify, select, and assess pertinent research which also involves collecting and analyzing data from the selected studies critically (Moher et al., 2009). Certain steps had to be skipped considering the duration and other limitations we had to deal with.

Overall, our approach aimed to provide a rigorous and systematic review of the literature, while also ensuring the feasibility of completing the study within a limited timeframe.

Data Collection

To identify relevant papers on data science and coding education, we conducted a search using two databases: Google Scholar and IEEE. The search was limited to papers published in English between 2010 and 2022. We used a combination of keyword searchers and operators, including "data science education", "coding education", "programming education", "data literacy", "data skills", "curriculum development", "industry needs", and "job market.

Duplicate papers were removed, and the remaining papers were assessed for relevance based on the research question and inclusion criteria. We identified a total of 16 papers from the screening process. Additionally, citation tracking was carried out in both the forward and backward directions to identify any additional relevant articles, which resulted in the inclusion of 2 papers in the final analysis. (Greenhalgh and Peacock, 2005)

Using Google Scholar and IEEE to search for papers on data science and coding education allowed for a comprehensive and focused approach to identifying relevant papers. By using keyword searchers and operators, the search was able to target specific concepts and themes related to the research question. Furthermore, citation tracking helped to identify additional relevant papers and strengthen the study's findings. (Booth, 2016).

Screening

After the search results were exported from the two databases, they were imported into a screening tool called Covidence which can significantly simplify the process of reviewing titles and abstracts for a JBI systematic review with an online platform that is available for free and also offers a subscription version that includes additional features. (Macdonald et al., 2016) This tool allowed for a more efficient and streamlined approach to screening the search results.

The first step of the screening process involved reviewing the titles and abstracts of the records to determine their relevance to the research question and inclusion criteria. If the title and abstract indicated that the record may meet the inclusion criteria, the record was selected for full-text screening.

During the full-text screening process, the inclusion and exclusion criteria were applied more rigorously. Records were reviewed in depth to determine if they met the criteria for inclusion in the systematic review. If a record did not meet the criteria, it was excluded from further consideration.

Throughout the screening process, reasons for exclusion were recorded in a separate document. This allowed for transparency and traceability in the screening process, ensuring that decisions were made systematically and consistently.

Overall, the use of a screening tool and a transparent process for recording exclusion reasons helped to ensure that the final selection of studies was rigorous and unbiased. This approach increases the reliability and validity of the systematic review and enhances the quality of the findings. 18 papers were initially used for the screening process and after the screening 09 papers become eligible for the next phase of extracting data.

Data Extraction

In this systematic review, data extraction was performed on a total of 09 selected papers. The following sections were extracted from each paper: Study ID, Study Design, Participants, Interventions/Exposures, Outcomes/Results, Limitations, and Other Notes.

Study ID: The unique identifier for each study was extracted, such as the author's name and publication year.

Study Design: The study design of each paper was extracted, such as randomized controlled trials, cohort studies, case-control studies, or cross-sectional studies.

Participants: The characteristics of the study participants were extracted, such as age, gender, ethnicity, and any relevant medical conditions.

Interventions/Exposures: The interventions or exposures used in each study were extracted, such as medication or treatment regimens, lifestyle modifications, or environmental exposures.

Outcomes/Results: The outcomes or results of each study were extracted, such as changes in disease incidence, symptom improvement, or adverse effects.

Limitations: Any limitations or potential biases in each study were extracted, such as small sample size, incomplete data, or confounding factors.

Based on the extracted data, a summary of the key findings and trends among the nine selected papers was compiled. For example, it may be noted that the majority of the studies were randomized controlled trials that focused on the effectiveness of coding knowledge for Data Science. Limitations may include small sample sizes or lack of diversity among study participants. Overall, the data extracted from each paper was used to inform the conclusions and recommendations presented in the systematic review. Below are the research papers that were used for data extraction along with the data that was extracted for data coding process.

1. *Skill Requirements in Big Data: A Content Analysis of Job Advertisements* (Gardiner et al., 2016) in their research paper explores the content analysis of job advertisements. For this, the participants were online job ads related to big data. The study used the big data disciplinary conceptual model of professional expertise, abilities, and skills (BDDSCM). The paper highlighted that many big data personnel create analytical systems, and many big data positions demand traditional development abilities. The results of our research also showed that businesses still place a high priority on soft skills for big data personnel. The paper provides the future scope that the current study may have been skewed toward technical positions,

underrepresenting big data-related positions that were more heavily weighted toward managerial decisions or simply statistical analysis.

2. *Integrating online meta-cognitive learning strategy and team regulation to develop students' programming skills, academic motivation, and refusal self-efficacy of Internet use in a cloud classroom* (Tsai1 et al., 2022) the research paper used a Quasi-experimental design for the university students were the participants. The study used the team management and online metacognitive learning plan integration model. The study explores the Students' programming design abilities and self-efficacy with Internet use in a virtual classroom both greatly improved after taking an online programming course. The results of this study did not reveal the anticipated impacts of meta-cognitive learning strategy (MCLS) on students' programming abilities, academic motivation, or resistance to self-efficacy of Internet usage. A modest sample size was used for the investigation. The study solely looked at how combining team regulation (TR) and online meta-cognitive learning strategy (MCLS) affected students' development of programming abilities, academic motivation, and resistance to self-efficacy of Internet usage.

3. *Game-Based Learning: Enhancing Student Experience, Knowledge Gain, and Usability in Higher Education Programming Courses* (Zhao et al., 2022) in their current study used over 100 students from 3 institutions took part in the study. 78, 10, and 34 students took part in the Variable game and 65, 10, and 30 students participated in the Loop game in DCU, NCI, and STUBA, respectively. The study makes the case that game-based learning is a useful intervention to improve the student experience, knowledge acquisition, and usability in higher programming education. A comprehensive assessment methodology which includes surveys, observations, and interviews were used. The pilot study's findings revealed that the majority of students, regardless of their geographic region, thought the games were both aesthetically pleasing and flawlessly functional. The majority of students enjoyed playing serious games because they believed that they improved their understanding of programming topics, piqued their interest in their classes, and helped them get higher grades. These outcomes show how effectively the suggested games accomplished their objectives. The study, which focuses on information retention, learner experience, and game usability, looked at how students responded to game-based learning. Based on their location, educational background, and the game they played, the study discovered that all pupils profited from the games, albeit to varied degrees. The study's limitations are mostly related to the technology that supports GBL, including the possibility of gadget addiction, poor posture, and decreased participant interest in other activities.

4. *Toward Knowledge Discovery Framework for Data Science Job Market in the United State* (Heidarysafa et al., 2022) the study highlighted that Python is the most popular language for work in data science, followed by R and Java. This shows that Python is not only the most popular language but also that, based on the trend and the gap, it would likely continue to be in demand. A data analyst primarily needs abilities for data retrieval (e.g., SQL, Excel, database), and visualization (e.g., Tableau, Power BI). Machine learning engineer job ads typically reflect a dominance of talents such as programming, machine learning, cloud, and big data technologies. On the other hand, an applicant for a position as a data scientist would require expertise in both of the aforementioned areas. Another significant finding from this study emphasizes the value of deep learning, cloud computing, and big data expertise for data scientists. Python appears more frequently than R, indicating that it is a more prevalent

language for occupations using data in the industry than R, which is another interesting finding in this context. The study was based on Quantitative Analysis. Using Knowledge Discovery Framework the research was limited only to the United States context.

5. *Problem-solving by 5–6 years old kindergarten children in a computer programming environment: A case study* (Fessakis et al., 2013) the study used the participant's 5-6-year-old kindergarten children. The study was based on an exploratory case study. The study found that children enjoyed engaging learning activities and had opportunities to develop mathematical concepts, problem-solving, and social skills. The intervention was designed as a part of the structured learning activities of the kindergarten which are teacher-guided and are conducted in a whole class social mode.

6. *School Perceptions of Coding Education in K-12: ALarge Scale Quantitative Study to Inform Innovative Practice* (Wong et al., 2015) the study was conducted on 42 primary and secondary schools in Hong Kong. The data was collected through a questionnaire. The article presents the findings of a study conducted in 42 Hong Kong schools. The findings indicate that there is a movement in the neighborhood schools to promote computer science education, with favorable views of both teaching and learning. However, the study also mentions issues with education that still need to be resolved, such as a lack of teacher training and an inconsistent curriculum. The limitation of the study is that the results may not apply to other locations as the study only examined schools in Hong Kong.

7. *Teaching Introductory Programming: a Quantitative Evaluation of Different Approaches* (Koulouri et al., 2013) the study used pair programming, the Media Computation approach, and Peer Instruction. The data was collected through Quantitative Analysis. The participants were a module leader, 2 supporting lecturers, 6 graduate teaching assistants, and the students in the class. The paper makes the following claims: it is important to rigorously control the process in a naturalistic setting; some objective measurements should be taken into consideration to assess the effects of the strategies employed; quantitative evidence has been produced from these cycles of data collection and analysis; and recommendations for the design of CS1 curricula have been framed as a result of these analyses. The study indicated that even while students liked it and found it simple to use, programming in Scratch did not significantly affect their ability to solve problems. According to the study, assessment criteria for instructional interventions should be rephrased as "indicators of learning."

8. *The Effects of Teaching Programming via Scratch on Problem-Solving Skills: a Discussion from Learners' Perspective* (Kalelioglu and Gülbahar, 2014) the study was conducted on 49 primary school students. Explanatory sequential mixed methods design is used to teach students from scratch. The study found that no significant increase in the mean of the factor of "self-confidence" in their problem-solving ability occurs.

9. *What is a Data Scientist? Analysis of core soft and technical competencies in job posting* (da Silveira et al., 2020) the study highlighted that Python, English, SQL, and good communication skills are the primary competencies needed for data science. According to Costa and Santos (2017), data scientists should have a profile that emphasizes their understanding of programming languages, databases, and machine learning. Google and Facebook, as quoted in Costa & Santos,(2017) have also drawn attention to the importance of analytical, mathematical, and statistical skills as well as the passion for using data to make choices and address issues. All of those qualities are frequently cited in the positions under

analysis, along with others. The fact that hard talents are listed in every opportunity, even if just briefly, indicates that the job postings value them the highest. The data itself has a very dynamic nature since new jobs are being posted and being fulfilled, which causes them to be deleted from the platform.

Data Synthesis

Data synthesis for this research was conducted using the Thematic Analysis process.

Firstly, the text from the outcomes and results section of the Data extraction was used to create codes by extracting significantly relevant texts to Research questions. These codes were specific phrases or sentences that were related to the research questions and provided insight into the data mentioned in Figure 2.

What are the factors and methods that contribute to the effectiveness of data science and coding education programs?	What are the main coding skills and requirements of the data Science job market?
Online Programming course significantly improved student's programming design skills.	Top Language for Data Science jobs are python following by R and Java.
Games helped them to understand better programming concepts;	Main skills required for Data Science are good communications, Team player, Problem solver, Python, English and SQL.
Students benefitted from the game, although to different degrees depending on their location, educational background, and the game played.	Hirers care more about the applicant's skills and knowledge than about his or hers formal academic education.
Children enjoyed the engaging learning activities and had opportunities to develop mathematical concepts, problem solving and social skills.	That many big data employess develop analytical system.

Figure 2. Data Coding Process.

Secondly, Figure 3 was created dividing the codes into two columns according to research questions. This helped to organize the codes and ensured that each code was assigned to the appropriate research question.

Outcomes/ Results	Codes
Skills associated with many big data jobs required traditional development skills, reflecting that many big data employees develop analytical systems. Furthermore, the Analysis also highlighted that soft skills are still highly valued by employees for big data employees.	Many big data employees develop analytical system, soft skills are still highly valued by employers for big data employees.
Online programming course significantly improve student's programming design skills and their refusal self-efficacy of internet use in a cloud classroom . However, the expected effects of meta-cognitive learning strategy (MCLS) on developing student's programming skills, academic motivation, and refusal self-efficacy of internet use were not found in this study.	Online programming course significantly improved students's programming design skills.

Figure 3. Organize Codes relevant to Research Questions.

Thirdly, themes were created by connecting and combining similar codes. This step involved reviewing the codes and identifying commonalities and patterns in the data. The codes were then grouped into themes based on these commonalities and patterns.

The final output of the data analysis was a set of themes that provided insight into the research questions. These themes were supported by the codes and provided a deeper understanding of the data.

Themes Derived from the Codes

Theme 1: Kids enjoy learning coding with fun activities and aesthetics

This theme highlights the importance of incorporating fun activities like Gaming and aesthetics in coding education to make it more enjoyable for children. It emphasizes that coding should be taught in a way that is both informative and entertaining to capture children's attention and interest. By doing so, children can learn to code more effectively and improve their problem-solving, critical thinking, and social skills. The theme also emphasizes that coding education should be designed to appeal to different learning styles, allowing children to learn at their own pace and in a way that suits them best.

Theme 2: More training and revised curriculum required for coding education

This theme emphasizes the need for more teacher training and a revised curriculum to ensure the successful delivery of coding education. It highlights the challenges faced by educators in teaching programming, such as the lack of standardized curriculum and teacher training. The theme advocates for a more comprehensive and standardized curriculum that provides students with the necessary skills and knowledge to succeed in the field of programming. It also emphasizes the importance of providing teachers with adequate training and support to ensure that they are equipped to teach coding effectively.

Theme 3: Programming Education can benefit children by improving their personal skills

This theme highlights the positive outcomes and benefits of programming education for children, such as an improvement in their personal skills. The theme emphasizes that coding education can help children develop problem-solving, critical thinking, creativity, and social skills. These skills are not only important for success in programming but also personal growth and development. The theme also highlights the potential for coding education to improve children's academic performance and self-confidence in their problem-solving abilities. Ultimately, programming education can provide children with valuable skills and knowledge that will benefit them in various aspects of their lives.

Theme 4: In the field of Data Science jobs, a candidate's soft skills and technical abilities such as coding hold more significance than their academic qualifications.

This theme highlights the changing landscape of hiring practices in the field of data science, where soft skills and technical abilities hold more weight than traditional academic qualifications. It focuses on the top programming languages preferred by employers in the field of data science, which are Python, R, and Java, and the primary skills required for data scientists, such as good communication, problem-solving, team playing, and proficiency in programming languages like Python and SQL. The theme also delves into the shifting priorities of hirers who now prioritize a candidate's skills and knowledge over formal academic education. Additionally, it touches upon the fact that many big data employees develop analytical systems, and employers still value soft skills, like communication and teamwork, in addition to technical abilities.

Conclusion

The findings from the themes discussed above reveal the significance of incorporating fun and engaging activities in coding education to improve children's learning outcomes. Additionally, it is crucial to provide teachers with the necessary training and a standardized curriculum to ensure the successful delivery of coding education.

Moreover, programming education has the potential to enhance children's personal skills, such as problem-solving, critical thinking, and creativity, that will benefit them in various aspects of their lives. This implies that programming education can serve as a valuable tool for personal growth and development. In the field of data science jobs, the significance of soft skills and technical abilities, particularly coding, has surpassed traditional academic qualifications. Employers prioritize the skills and knowledge of candidates over formal academic education, and programming languages like Python, R, and Java have become essential for data scientists. Good communication, problem-solving, and team-playing skills are also vital for success in this field. Thus, to succeed in the rapidly evolving field of data science and programming, it is essential to focus on developing technical abilities, like coding, as well as soft skills. Overall, programming education offers a range of benefits and outcomes, making it an essential tool for personal and professional growth in today's world.

The findings on coding education and the field of data science are closely linked. As children are introduced to coding education at a young age, they can develop fundamental skills in problem-solving, critical thinking, and programming languages like Python, which are essential for data science jobs. The emphasis on incorporating fun activities and appealing to different learning styles in coding education can also foster creativity and communication skills, which are valuable assets in the data science job market. The need for more teacher training and a revised curriculum in coding education aligns with the demand for a comprehensive skill set in the field of data science, as employers prioritize candidates who possess both technical abilities and soft skills like communication and teamwork. Thus, investing in coding education can not only benefit children but also prepare them for a promising career path in the ever-growing field of data science.

This research was conducted under some limitations like time & number of reviewers available, therefore some steps had to skip and proceed in a rapid review manner. On a future research note, this can be done by expanding inclusion and exclusion criteria further by gathering papers from more wide time frame and also gathering papers from different contexts and languages. Also, the methodology used in this systematic review could be applied to various research questions in the field of data science and coding education, with potential future research directions focusing on specific target populations or the skills and knowledge required for data science and coding careers.

References

Ari, F., Arslan-Ari, I., & Vasconcelos, L. (2022). Early childhood preservice teachers' perceptions of computer science, gender stereotypes, and coding in early childhood education. *TechTrends*, *66*(3), 539-546.

Braun, V., & Clarke, V. (2016). (Mis) conceptualising themes, thematic analysis, and other problems with Fugard and Potts'(2015) sample-size tool for thematic analysis. *International Journal of social research methodology*, *19*(6), 739-743.

Booth, A. Searching for qualitative research for inclusion in systematic reviews: a structured methodological review. Syst Rev. 2016; 5: 74.

Davenport, T. H., & Patil, D. J. (2012). Data scientist. *Harvard business review*, *90*(5), 70-76.

Duda, M., Sovacool, K. L., Farzaneh, N., Nguyen, V. K., Haynes, S. E., Falk, H., ... & Wolford, B. N. (2021). Teaching Python for Data Science: Collaborative development of a modular & interactive curriculum. *The Journal of open source education*, *4*(46).

da Silveira, C. C., Marcolin, C. B., da Silva, M., & Domingos, J. C. (2020). What is a Data Scientist? Analysis of core soft and technical competencies in job postings. *Revista Inovação, Projetos e Tecnologias*, *8*(1), 25-39.

Emery, N. C., Crispo, E., Supp, S. R., Farrell, K. J., Kerkhoff, A. J., Bledsoe, E. K., ... & Aiello-Lammens, M. E. (2021). Data science in undergraduate life science education: a need for instructor skills training. *BioScience*, *71*(12), 1274-1287.

Fessakis, G., Gouli, E., & Mavroudi, E. (2013). Problem solving by 5–6 years old kindergarten children in a computer programming environment: A case study. *Computers & Education*, *63*, 87-97.

Fischer, M., Gerdtham, U. G., Heckley, G., Karlsson, M., Kjellsson, G., & Nilsson, T. (2021). Education and health: long-run effects of peers, tracking and years. *Economic Policy*, *36*(105), 3-49.

Feijóo-García, P. G., McNamara, K., & Stuart, J. (2020, March). The Effects of Native Language on Block-Based Programming Introduction: A Work in Progress with Hispanic Population. In *2020 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)* (Vol. 1, pp. 1-2). IEEE.

Glen, S. (2021). *Data Science Job Market Shrinking? Not So Fast.* Retrieved from https://www.datasciencecentral.com/.

Greenhalgh, T., & Peacock, R. (2005). Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *Bmj*, *331*(7524), 1064-1065.

Gardiner, A., Aasheim, C., Rutner, P., & Williams, S. (2018). Skill requirements in big data: A content analysis of job advertisements. *Journal of Computer Information Systems*, *58*(4), 374-384.

Heidarysafa, M., Kowsari, K., Bashiri, M., & Brown, D. E. (2022). Towards a Knowledge Discovery Framework for Data Science Job Market in the United States. In *Proceedings of the Future Technologies Conference (FTC) 2021, Volume 1* (pp. 875-887). Springer International Publishing.

Hundhausen, C. D., Agrawal, A., & Agarwal, P. (2013). Talking about code: Integrating pedagogical code reviews into early computing courses. *ACM Transactions on Computing Education (TOCE)*, *13*(3), 1-28.

Horton, N. J., & Hardin, J. S. (2021). Integrating computing in the statistics and data science curriculum: Creative structures, novel skills and habits, and ways to teach computational thinking. *Journal of Statistics and Data Science Education*, *29*(sup1), S1-S3.

Jindal, P., & Tej, F. (2016). Being an educator in private vs. public school system: A comparative study on job contentment. *IRA International Journal of Education and Multidisciplinary Studies*, *3*(3), 309-319.

Kumari, S., Jindal, P., & Chand, P. K. (2022). Social Media Networking: As Facilitator of Acculturation. *ECS Transactions*, *107*(1), 9381.

Koulouri, T., Lauria, S., & Macredie, R. D. (2014). Teaching introductory programming: A quantitative evaluation of different approaches. *ACM Transactions on Computing Education (TOCE)*, *14*(4), 1-28.

Kalelioglu, F., & Gülbahar, Y. (2014). The Effects of Teaching Programming via Scratch on Problem Solving Skills: A Discussion from Learners' Perspective. *Informatics in education*, *13*(1), 33-50.

Karaoglu, G., Hargittai, E., & Nguyen, M. H. (2022). Inequality in online job searching in the age of social media. *Information, Communication & Society*, *25*(12), 1826-1844.

Karahan, E., BİLİCİ, S. C., & Ayçin, Ü. N. A. L. (2015). Integration of media design processes in science, technology, engineering, and mathematics (STEM) education. *Eurasian Journal of Educational Research*, *15*(60), 221-240.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute.

Miller, J. (2019). STEM education in the primary years to support mathematical thinking: Using coding to identify mathematical structures and patterns. *Zdm*, *51*(6), 915-927.

Mikroyannidis, A., Domingue, J., Bachler, M., & Quick, K. (2018, October). Smart blockchain badges for data science education. In *2018 IEEE Frontiers in Education Conference (FIE)* (pp. 1-5). IEEE.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, *98*, 261-276.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group*, T. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, *151*(4), 264-269.

Macdonald, M., Misener, R. M., Weeks, L., & Helwig, M. (2016). Covidence vs Excel for the title and abstract review stage of a systematic review. *JBI Evidence Implementation*, *14*(4), 200-201.

Oesterreich, T. D., Teuteberg, F., Bensberg, F., & Buscher, G. (2019). The controlling profession in the digital age: Understanding the impact of digitisation on the controller's job roles, skills and competences. *International journal of accounting information systems*, *35*(C).

Prato, S. C. (2017). Beyond the computer age: A best practices intro for implementing library coding programs. *Children and Libraries*, *15*(1), 19-21.

Pournaras, E. (2017). Cross-disciplinary higher education of data science–beyond the computer science student. *Data Science*, *1*(1-2), 101-117.

Tsai, C. W., Lee, L. Y., Cheng, Y. P., Lin, C. H., Hung, M. L., & Lin, J. W. (2022). Integrating online meta-cognitive learning strategy and team regulation to develop students' programming skills, academic motivation, and refusal self-efficacy of Internet use in a cloud classroom. *Universal Access in the Information Society*, 1-16.

Wong, G. K., Cheung, H. Y., Ching, E. C., & Huen, J. M. (2015, December). School perceptions of coding education in K-12: A large scale quantitative study to inform innovative practices. In *2015 IEEE international conference on teaching, Assessment, and learning for engineering (TALE)* (pp. 5-10). IEEE.

Zhao, D., Muntean, C. H., Chis, A. E., Rozinaj, G., & Muntean, G. M. (2022). Game-based learning: enhancing student experience, knowledge gain, and usability in higher education programming courses. *IEEE Transactions on Education*, *65*(4), 502-513.