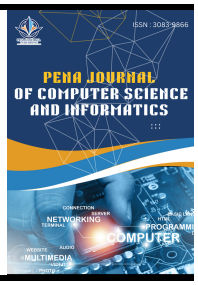
PENA CENDEKIA PUBLISHING
NURTURING IDEAS, ENRICHING KNOWLEDGE

Pena Journal of Computer Science and Informatics

Journal homepage:
https://penacendekia.com.my/index.php/pjcsi/index
ISSN: 3083-9866



A Comparative Analysis of Database Topic Difficulty in Relation to Web Development Readiness

Nur Dalila Abdullah[1,*], Azra Mohammad Amirruddin[1] , Nurul 'Afifah Razali[1]

[1] Department of Information Technology and Communication, Politeknik Muadzam Shah, Pahang, Malaysia

ARTICLE INFO

Article history:
Received 5 October 2025
Received in revised form 10 November 2025
Accepted 15 November 2025
Available online 22 November 2025

Keywords:
Database; topic difficulty; diagnostic assessment; web development readiness

ABSTRACT

This study investigates topic-level difficulties in the Database Systems course and their implications for Web Development readiness among diploma students. A total of 129 students from four classes 2A, 2B, 2C, and 2D enrolled in the Diploma in Information Technology programme participated in this study. Final examination scores were mapped across five topics which are The Fundamentals of Database Management System, Relational Data Models, Normalisation and Entity-Relationship Diagrams (ERD), Structured Query Language (SQL), and Transaction Management. Descriptive statistics were used to identify topic-level performance, while Spearman's rank-order correlation examined relationships between topic mastery and overall achievement. The descriptive analysis shows that Topic 1 with mean score of 0.87 and a standard deviation is 0.21, was the most understandable topic indicating that students generally had a firm grasp of foundational concepts, while Topic 2 with mean score is 0.63 and standard deviation is 0.22, was the most difficult, confirming relational modelling as a persistent area of weakness. Results further revealed that SQL demonstrated the strongest correlation with overall scores ($\rho = 0.771$, $p < 0.001$), whereas Transaction Management contributed the least ($\rho = 0.328$, $p < 0.001$). These findings underscore the importance of diagnostic assessments for identifying recurring weaknesses and guiding targeted pedagogical strategies. Strengthening competence in relational modelling and SQL at this stage is crucial, as these skills underpin effective server-side programming and database integration in the Semester 4 Web Development course.

1. Introduction

Mastering databases is crucial in computing education as it establishes the basis for creating dynamic, data-oriented applications. In higher education, assessing particular issues in database courses provides essential understanding of students' readiness for advanced topics, like web development. Previous studies highlight that challenges in aspects such as Entity Relationship

* *Corresponding author.*
E-mail address: dalila@pms.edu.my

https://doi.org/10.37934/pjcsi.3.1.110

Diagrams (ERD), normalization and Structured Query Language (SQL) frequently obstruct advancement into applied programming subjects where database integration is essential [1].

The Diploma in Information Technology (DIT) program is structured to facilitate learning sequentially. In the second semester, students must take the Database course as it serves as a prerequisite for the Web Development course in the fourth semester. The students will apply the database concepts they have learned with Hypertext Pre-processor (PHP) and MySQL integration for the web development course. Therefore, it is compulsory to pass the database course before moving forward as this ensures that students develop essential skills in schema design and query execution, which are important for back-end web programming.

Despite this intentional curriculum design, differences in topic mastery often appear between classes, even within the same group of students. Some topics consistently present more difficulties, resulting in weaker foundations that may affect students ability to move into web development. Identifying these tough topics is essential for instructors to improve teaching methods, manage time better, and prepare students for future coursework.

Existing research consistently shows that students encounter uneven experience levels of difficulty across database topics. Basic concepts, such as defining databases or identifying Database Management System (DBMS) components, are usually easier to grasp, whereas advanced areas including relational modelling, normalization, SQL joins, and transaction management, tend to be remain challenging and often lead to lower performance [2]. These difficulties affect success in the database course itself and compromise readiness for web development, where database integration skills are essential. Recent studies have therefore emphasized diagnostic, topic-level analyses as valuable tools for pinpointing specific weaknesses and designing targeted instructional interventions [3]. Aligning teaching with the results of diagnostic evaluations allows educators to more effectively tackle specific gaps in knowledge, enhance students understanding of database concepts and ultimately promote greater achievement in web development and other advanced topics.

Analyzing examination performance at the topic level is an effective way to identify areas where students face difficulties and to develop more efficient teaching strategies. This research evaluates the final exam scores from four groups in the same cohort to examine the challenges of different subjects in database education. It utilizes descriptive statistics and item-level comparisons to identify weak student performance areas. This information will help guide targeted weaknesses, ensuring that students are better prepared for subsequent courses, such as Web Development.

1.1 Research Objectives

The following objectives guide this study:

i. To analyze the relative difficulty of database topics among students in four classes within the same cohort.
ii. To compare student performance across topics to identify recurring areas of weakness.
iii. To evaluate how variations in database topic mastery may influence students' readiness for the Web Development subject.

2. Literature Review

2.1 Curriculum Design and Prerequisites in IT Programmes

Diploma in Information Technology (DIT) programmes in Malaysia are designed with a structured sequence that builds progressively from foundational computing skills to applied system development. In many polytechnics, the database subject is introduced in the early semesters to

provide students with essential data organisation and management knowledge. This placement is intentional, as subsequent courses particularly those involving dynamic websites or server-side programming require students to possess already the ability to design, manipulate, and query data effectively. The prerequisite ensures that students who progress to more advanced modules such as web application development are equipped with the necessary competencies to integrate backend databases with front-end interfaces.

Academic institutions worldwide observe similar curricular patterns. For instance, Lakeshore Technical College's Web and Software Developer program integrates database administration early in its curriculum to support later application development [4]. Likewise, at the University of Southern California (USC), web development courses are sequenced sequentially where students progress from Intro to Web Development, through front-end, then back-end, culminating in a web application project each step requiring successful completion of the previous [5].

This criterion exists in private institutions such as Universiti Poly-Tech Malaysia, where the curriculum incorporates database classes prior to web programming [6]. By incorporating this foundational framework, both public and private institutions guarantee academic success in aiding students to not only excel in coding but also to grasp the crucial data infrastructure that underpins web systems. This method of balancing foundational and applied topics mirrors wider patterns in Malaysian Information Technology education, which aims to generate graduates that align with industry demands for database-oriented application development. These organized frameworks highlight that expertise in databases is a crucial necessity for students success in web development and the attainment of skills relevant to and aligned with future industry needs.

2.2 Core Topics in Database Course

The Database System course provides students with an extensive overview of data organization and management. It starts with the basics of DBMS, in which learners grasp essential principles, the structure of databases, and the advantages of utilizing databases over conventional file systems. The course subsequently examines the Relational Data Model, highlighting how data can be represented through tables, attributes, and connections. This part covers the use of keys and constraints to ensure data integrity and the course moves on to Normalization and ERD, where learners will use abstract thinking to organize data in a systematic way. This phase seeks to minimize duplication and confirm precision in data modelling.

The development advances with SQL in which students gain hands-on expertise to accurately define, manage, and query data. The course then addresses Transaction Management, presenting fundamental ideas like ACID properties, concurrency control, and methods for recovery. These subjects form a systematic educational route that improves students skills in database management and provides them with crucial insights for web development. By fully equipped knowledge from this domain, students are likely to be prepared by combining databases with server-side coding, developing dynamic, secure, and dependable web applications [7,8].

2.3 Database as a Foundation for Web Development

Databases are fundamental to most modern web applications, and understanding database concepts is crucial for effective web development. Learning to design, normalize, and query databases provides essential skills for managing dynamic web content and ensuring application reliability. Research has shown that students who struggle with foundational database courses often face difficulties in later web development subjects, especially when integrating server-side scripting

with relational data models. A key challenge in this progression is the uneven difficulty across various subtopics within databases. As noted by Abdelaziz *et al.,* [9], learners typically find it easier to grasp introductory concepts, such as defining databases and identifying the components of DBMS. However, they often struggle with more advanced topics like relational database, joins clause, and transaction management.

Recent studies highlight the significance of database knowledge as a predictor of success in web development. Miedema *et al.,* [10] demonstrated that a solid understanding of normalization and entity-relationship modelling directly correlates with student's ability to build scalable and functional web applications. Additionally, Radović *et al.,* [11] observed that students who engage in problem-based and gamified approaches to learning databases perform better in later web programming tasks. This suggests mastering fundamental database skills can accelerate readiness for more advanced development courses.

Furthermore, topic-level diagnostic assessments offer instructors valuable insights into student weaknesses. By identifying specific database topics where students struggle, educators can design targeted remedial strategies that improve database skills and prepare students for web development. This approach aligns with broader educational recommendations that support structured curricular pathways, where learning about databases is a foundation for courses in server-side programming, application frameworks and full-stack development.

The literature collectively confirms that mastering databases is not just an isolated academic objective but a crucial foundation for successful web development. Therefore, it is essential to implement targeted teaching strategies to ensure strong performance in database courses. This approach will help produce graduates skilled in designing, developing, and maintaining robust web applications.

3. Methodology

3.1 Research Design

This study adopted a quantitative, comparative and correlational research design to examine the difficulty of various topics in a Database Systems course, and whether this difficulty relates to students readiness for web development. This design is suitable as it enables systematic analysis of students' performance across different database topics, allows for class comparisons, and helps identify correlations between mastery of issues and overall achievement. The study involved descriptive statistics and inferential testing to provide diagnostic insights and predictive implications.

3.2 Participant and Examination Structure

The participants of this study were students enrolled in the Diploma in Information Technology (DIT) programme and a total of 129 students from four classes were included where class 2A, n = 33, class 2B, n = 31, class 2C, n = 33 and class 2D, n = 32. All students are currently in their second semester and enrolled in the Database Systems course. This course is a compulsory subject and a prerequisite for Web Development in Semester 4. This means that students must pass the Database Systems course in order to progress.

3.3 Instrument and Data Source

The instrument used in this study was the final examination paper for the Database Systems course. This exam was designed based on the official course syllabus and covered five major topics

which are Fundamentals of DBMS, Relational Data Models, Normalization and ERD, SQL, and Transaction Management. It consisted of 30 multiple-choice questions, totalling 45 marks, with the marks distributed proportionately to reflect the weightage of each topic.

Each student's scores were calculated based on individual topics, and the percentage achievement was determined relative to the maximum possible score for each topic. Consequently, the final dataset included topic-level scores, overall course performance, and aggregated class statistics. This data served as the primary tool for analyzing topic difficulty, inter-topic relationships, and predictors of overall performance. Table 1 summarizes the examination structure, showing the marks distribution and number of questions across five grouped topic areas. This structure served as the basis for analyzing student performance at the topic level

Table 1
Examination structure for database course

Topic	Marks	Number of Questions
Topic 1	4.5	3
Topic 2	7.5	5
Topic 3	12	8
Topic 4	16.5	11
Topic 5	4.5	3
Overal	45	30

This design reflects a deliberate content weighting, with Topic 4 receiving the highest allocation of 16.5 marks and 11 questions, followed by Topic 3 with 12 marks and eight questions, emphasizing the importance of advanced concepts. In contrast, Topics 1 and 5 were assessed through fewer items with 4.5 marks and three questions each, primarily covering introductory or supplementary concepts. The uniform application of this structure ensured comparability of results across all four classes, thereby enabling meaningful analysis of topic-specific difficulty patterns.

3.4 Data Collection Procedures

The examination papers were marked using a standardized marking scheme, and the individual student scores were compiled into an Excel database. This dataset was then imported into IBM SPSS Statistics (Version 31) for analysis. Scores were categorized by the five topics, allowing for a detailed topic-level diagnostic analysis. The dataset was organized by classes of 2A, 2B, 2C and 2D to facilitate within-class and cross-class comparisons. The comparative designs are practically used in educational research to examine topic difficulty variations and to notify earlier the teaching strategies [12,13].

3.5 Ethical Considerations

Ethical principles were applied throughout this study. The approval for student examination data was obtained from the examination department and all data were anonymised before analysis to protect the identity and confidentiality of students. No individual student was identifiable and mentionable in reporting results and findings were presented only in accumulated form. The study did not involve any intervention or manipulation of the marks beyond regular academic assessment. Therefore, it posed no risk of harm to participants. Participation was viewed as implicit because the analysis depended only on available examination data. The study followed institutional regulations and conformed to ethical principles for educational research.

4. Results and Discussion

4.1 Descriptive Analysis for Topic Difficulties

The descriptive analysis of examination results across the four classes of 2A, 2B, 2C, and 2D revealed variations in students mastery of the five database topics. Table 2 summarizes the mean scores and standard deviations for each topic by class. The mean (M) indicates the average level of performance for each topic, while the standard deviation (SD) reflects the degree of score dispersion around the mean. A lower SD suggests that students' performance is more consistent, with most students scoring close to the mean. In contrast, a higher SD indicates greater variability, implying that some students performed well while others struggled significantly.

Table 2
Mean and standard deviation

Class	Topic 1		Topic 2		Topic 3		Topic 4		Topic 5	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
2A	1.00	0.00	0.64	0.20	0.74	0.18	0.72	0.18	0.53	0.19
2B	0.80	0.25	0.62	0.24	0.81	0.13	0.75	0.14	0.76	0.23
2C	0.82	0.25	0.60	0.25	0.76	0.16	0.71	0.15	0.85	0.22
2D	0.83	0.19	0.65	0.18	0.81	0.14	0.69	0.15	0.90	0.17
Overall	0.87	0.21	0.63	0.22	0.78	0.16	0.72	0.16	0.76	0.25

Class 2A demonstrated the strongest mastery in Topic 1 at the class level with a perfect mean score of 1.00 and SD of 0.00, indicating all students successfully addressed the related questions. However, this group showed the lowest performance in Topic 5 with a mean of 0.53 and an SD of 0.19, suggesting significant difficulties in applying more advanced database concepts. In contrast, Class 2B showed relatively balanced performance, with the highest score in Topic 3 with a mean score of 0.81 and SD of 0.13 and weaker outcomes in Topic 2 with a mean score of 0.62 and SD of 0.24.

Class 2C exhibited similar difficulties with Topic 2, with a mean score of 0.60 and an SD of 0.25, reinforcing the trend that relational modelling was the most challenging area across groups. On the other hand, this class performed better in Topic 5 with a mean score of 0.85 and an SD of 0.22, suggesting stronger conceptual understanding of transaction-related principles compared to 2A. Class 2D also reflected this strength in Topic 5 with the highest mean among all groups, with a mean score of 0.90 and SD of 0.17, although their performance in Topic 4 was slightly lower, with a mean score of 0.69 and SD of 0.15.

Figures 1 illustrate the comparison of topic difficulties across classes, and Figure 2 illustrates the overall topic difficulties. When aggregated across all four classes of 129 participants, the overall analysis shows that Topic 1, with a mean score of 0.87 and an SD of 0.21, was the most understandable, indicating that students generally had a firm grasp of database fundamentals. In contrast, Topic 2, with a mean score of 0.63 and an SD of 0.22, was the most difficult across the cohort, confirming that relational modelling concepts remain a persistent area of struggle. Topics 3, 4, and 5 displayed intermediate difficulty levels, suggesting that while students progressed in normalization, ERD, and transaction management, these topics still required targeted reinforcement.

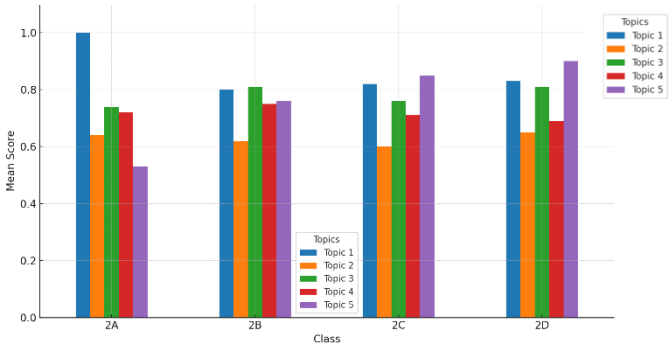


Fig. 1. Comparison of topic difficulties across classes

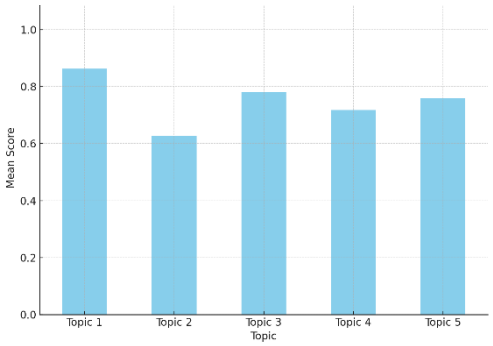


Fig. 2. Overall topic difficulties

These findings align with prior research that highlights how foundational database concepts are usually more accessible, while relational modelling and advanced query construction remain problematic for learners. In particular, difficulties in Topic 2 may be attributed to the abstract reasoning required for translating real-world scenarios into relational models, a skill that students often take longer to internalize. The variation observed between classes suggests that instructional strategies, learning dynamics, or prior knowledge may have influenced outcomes, reinforcing the importance of topic-level diagnostic assessment to guide targeted pedagogical intervention.

4.2 Correlation Analysis Between Topic Performance and Overall Achievement

A normality test, the Shapiro–Wilk test, was conducted for each topic and overall marks. The results in Table 3 indicate that all variables violated the normality assumption, with $p < 0.05$ across all topics and total scores. This suggests that the data were not normally distributed. One possible reason is that the weightage of marks differed across topics, as the examination was based on the official syllabus structure. These weightages were predetermined by the Final Examination Instrument Specification Table (FEIST) framework when the examination questions were constructed. As such, the unequal distribution of marks across topics likely influenced the non-normal distribution of the data. Since the normality assumption was violated, the non-parametric Spearman's rho correlation was employed instead of Pearson's correlation. The Spearman test is robust for ordinal or non-normal continuous data and has been widely recommended for educational assessment contexts [14-16].

Table 3
Tests of normality for topic scores and total marks

Variable	Shapiro-Wilk Statistic (W)	Sig. (p-value)
Topic 1	0.641	< .001
Topic 2	0.915	< .001
Topic 3	0.911	< .001
Topic 4	0.949	< .001
Topic 5	0.787	< .001
Overal	0.973	0.012

The Spearman correlation results are shown in Figure 3. Topic 4 demonstrated the strongest positive correlation with total marks, $\rho = 0.771$ and $p < 0.001$, suggesting that performance in this topic is the most influential predictor of overall achievement. In contrast, Topic 5 showed the weakest correlation, $\rho = 0.328$ and $p < 0.001$, implying that success in this topic contributes less directly to overall exam performance. Topics 2 with $\rho = 0.557$ and $p < 0.001$ and Topic 3 with $\rho = 0.568$ and $p <$

0.001 also showed moderate correlations, highlighting their role as core components of student mastery. Topic 1 had the lowest but still significant positive relationship with overall marks with ρ = 0.333 and p < 0.001.

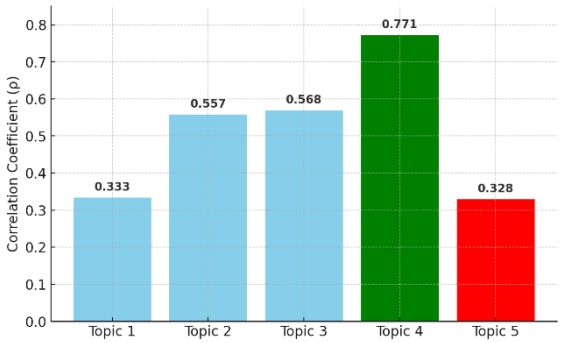


Fig. 3. Comparison of topic difficulties across classes

Overall, these findings suggest that students mastery of relational modelling and complex queries in Topic 4 plays a central role in determining total exam performance. In contrast, more isolated or less integrative topics like Topic 5 contribute less substantially. This reinforces the need for targeted instructional support in areas that heavily impact overall achievement, aligning with prior research on the diagnostic value of topic-level assessment in computing education.

4.3 Implications for Web Development Readiness

The results of this study show some interesting trends in student performance. While students tended to grasp the foundational topics in databases quite well, they faced significant challenges with more complex areas like relational modelling, SQL operations, and normalization. These struggles aren't unique; they echo findings from earlier studies that indicate learners often find it challenging to translate real-world situations into relational structures and to create more advanced queries. This is more than just a matter of academic difficulty, these gaps in understanding have important implications for students, especially as they prepare for the Web Development course in Semester 4. Addressing these weaknesses will be crucial for their overall success in the program.

Web development not only designing the user interface but also combines with the server-side scripting such as PHP and the database, MySQL. The Create, Read, Update and Delete (CRUD) operations is the common features in web development to adapt dynamic web application. It is essential for students to be strong foundation in Database Systems in order to success and understanding on how to incorporate in web development. Student who has not mastered the relational modelling or SQL most probably will facing challenges when implementing the server-side functionalities. This could complicate linking databases to user-friendly front-end interfaces and maintaining data consistency throughout their applications.

These findings highlight an important finding where students who are skilled in SQL and relational modelling can perform and are better prepared for the Web Development course. Early diagnostic assessment might help lecturers plan ahead to overcome the difficulties in understanding the integration of the database into the Web Development course. The differences between various classes suggest that some groups of students might start Semester 4 with a stronger foundation than others. Lecturers can explore the gamification technique and focus on the identified topics where students performed poorly such as SQL practice sessions, practical modelling exercises, and collaborative tasks. This approach can cater the knowledge gaps by focusing on these identified problematic topics during the Database Systems course. It can enhance students learning

experiences, allowing them to tackle server-side scripting and database integration more confidently. Strengthening their grasp of database concepts at this stage improves their immediate DBMS results and sets them up for success in Web Development. This ultimately ensures a smoother transition throughout the Diploma in Information Technology curriculum.

5. Conclusion

This study looked into how students performed in Database Management Systems across four different classes, highlighting their strengths and challenges. Overall, students showed a solid grasp of the foundational concepts, but there were notable struggles with relational modelling, normalization, and SQL operations. These areas were significant indicators of how well they would do on their exams. If these weaknesses aren't addressed, they could hinder students as they prepare for the Semester 4 Web Development course, where skills in server-side scripting and database integration are essential.

The findings highlight how valuable topic-level diagnostic assessments can be for identifying students needing extra help early on. By using focused remediation strategies like hands-on SQL practice labs, engaging problem-based learning activities, and collaborative modelling projects educational institutions can enhance students understanding of essential database skills. This leads to better immediate outcomes in database management systems and helps students transition more smoothly into web development courses. Future research could build on this work by examining student performance over multiple semesters or by investigating how different teaching methods, such as gamification or adaptive learning platforms, impact students' readiness for more advanced programming courses.

Acknowledgement

This research was not funded by any grant.

References

[1] Slonka, Kevin J., and Neelima Bhatnagar. "Teaching database normalization: do prerequisites matter." *Issues in Information Systems* 25, no. 1 (2024): 128-135. https://doi.org/10.48009/1_iis_2024_111

[2] Miedema, Daphne, George Fletcher, and Efthimia Aivaloglou. "Expert perspectives on student errors in SQL." *ACM Transactions on Computing Education* 23, no. 1 (2022): 1-28. https://doi.org/10.1145/3551392

[3] Roy, Sudeepa, Amir Gilad, Yihao Hu, Hanze Meng, Zhengjie Miao, Kristin Stephens-Martinez, and Jun Yang. "How Database Theory Helps Teach Relational Queries in Database Education (Invited Talk)." In 27th International Conference on Database Theory (ICDT 2024), pp. 2-1. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2024. https://doi.org/10.4230/LIPIcs.ICDT.2024.2

[4] Muraski, John Michael, and Jakob Iversen. "Growing computer science and information technology education in K-12: Industry demand and ecosystem support." *Journal of the Midwest Association for Information Systems (JMWAIS)* 2022, no. 2 (2022): 2. https://doi.org/10.17705/3jmwa.000075

[5] USC Viterbi School of Engineering. "Web Development Curriculum Structure." USC TAC. 2025. https://tac.usc.edu/academics/web-development/

[6] Universiti Poly-Tech Malaysia. *"List of Courses in Diploma in Computer Application Development."* 2024. https://www.uptm.edu.my/index.php/bit-in-computer-application-development

[7] Elalouf, Amir, Arik Edelman, Dafna Sever, Shaked Cohen, Reut Ovadia, Omer Agami, and Yulia Shayhet. "Students' perception and performance regarding structured query language through online and face-to-face learning." In *Frontiers in Education*, vol. 7, p. 935997. Frontiers Media SA, 2022. https://doi.org/10.3389/feduc.2022.935997

[8] Zhang, Shuwei. "Full-Stack Web Development: Building Comprehensive Applications." *Journal of Computing and Electronic Information Management* 13, no. 2 (2024). https://doi.org/10.54097/1aut7ey3

[9] Abdelaziz, Noha, Mohammed F. Farghally, Mostafa Mohammed, and Taysir Soliman. "Identifying and addressing common undergraduate database design misconceptions." In *Frontiers in Education*, vol. 10, p. 1554089. Frontiers, 2025. https://doi.org/10.3389/feduc.2025.1554089

[10] Miedema, Daphne, George Fletcher, and Efthimia Aivaloglou. "Expert perspectives on student errors in SQL." *ACM Transactions on Computing Education* 23, no. 1 (2022): 1-28. https://doi.org/10.1145/3551392
[11] Radović, Maja, Katarina Karić, and Đurađ Milošević. "Learning Challenges and Performance in the Databases Course." In *10th International Scientific Conference Technics, Informatics and Education-TIE 2024*. Faculty of Technical Sciences Čačak, University of Kragujevac, 2024. https://doi.org/10.46793/TIE24.402R
[12] Creswell, John W., and J. David Creswell. *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications, 2017.
[13] Creswell, John W. "Research designs. Qualitative, quantitative, and mixed methods approaches." (2009).
[14] Ghasemi, Asghar, and Saleh Zahediasl. "Normality tests for statistical analysis: a guide for non-statisticians." *International journal of endocrinology and metabolism* 10, no. 2 (2012): 486. https://doi.org/10.5812/ijem.3505
[15] Akoglu, Haldun. "User's guide to correlation coefficients." *Turkish journal of emergency medicine* 18, no. 3 (2018): 91-93. https://doi.org/10.1016/j.tjem.2018.08.001
[16] Statistics, Laerd. "Spearman's rank-order correlation." Laerd Statistics (2013): 33.