



Preliminary Study on SQLearn : Sql Query Learning Tool Using Close Ended Drag and Drop Approach

Putra Prima Arhandi¹, Annisa Taufika Firdausi², Banni Satria Andoko³, Muhammad Shulhan Khairy⁴, Muhammad Ilham Adhim⁵, and Ahmad Afif Supianto⁶

^{1,2,3,4,5} State Polytechnic of Malang, Malang 65141, Indonesia

⁶ National Research and Innovation Agency (BRIN), Central Jakarta 10340, Indonesia

⁶ Norwegian University of Science and Technology, Ålesund 6025, Norway

¹putraprima@polinema.ac.id

²annisa.taufika@polinema.ac.id

³ando@polinema.ac.id

⁴khairy@polinema.ac.id

⁵ilhamm179@gmail.com

⁶ahma063@brin.go.id

Abstract. Structured Query Language (SQL) is an industry standard for writing query. Writing SQL query is one of main competencies that must be possessed when learning database. Writing an SQL query could be challenging for novice students. SQLearn is a web-based application that give several cases study and questions. Our hypothesis is learning by using SQLearn will improve the students' understanding in SQL query. SQLearn uses close ended questions approach. Close ended questions are used to minimize possibility of syntax error. The lecturer will create the open-ended questions with a predefined answers and set of draggable SQL blocks. Student can learn while answer the questions by dragging provided SQL blocks into the correct position, when students finish dragging SQL Blocks, SQLearn will check their answer automatically. Twenty-eight sophomore students from Information Technology department are involved in the experiment. Pre-test and post-test are used to measure students' understanding before using and after using SQLearn. T-test analysis result shows that there is a significant difference between pre-test and post-test score where the post-test score is overachieved the pre-test score. Based on the early result analysis, a promising result occurred, where the use of SQLearn could improve the students' test score.

Keywords: SQL Query, Close Ended, Drag and Drop, Learning.

1 Introduction

The ability to master databases is one of the most important skills at this time. Databases are implemented in many fields not only in computer science and business. In the fields of computer science and software engineering, databases have become compulsory subjects [1], [2]. In the field of statistics, databases have become one of the technologies

that have been adopted, especially on the topic of computational statistics [3], database also used to storing, retrieving and searching dataset on biology related research [4].

Several research have been conducted on the competencies that a student must acquire when learning database. One of them is a standard created by the IEEE and the Association for Computing Machinery (ACM). This standard formulates the main competencies and supporting competencies in creating a learning material about databases[1]. According to the IEEE and ACM standards as well as several other studies, it can be concluded that several competencies that must be possessed include relational databases, SQL Query, conceptual modeling, normalization, and relational algebra [5] [6].

Writing SQL queries is an important competency but in practice, writing an SQL query has its challenges. SQL queries are written in declarative and set, this is very different from what is generally learned in computer programming where the programming language is written in an imperative approach[7]. Writing a SQL query considered as an activity that increase the cognitive load experienced by the query writer [8]. Students also often experience many errors due to this high cognitive load, because they have to remember the database schema, table names, and column names. Other errors also occur due to a lack of understanding of students regarding basic SQL syntax and the concept of relations in databases, especially when grouping and joining [9]. [10].

Despite having a different writing style from other programming languages, relational databases and SQL still dominate database implementation in industry and academia [11]. The query syntax that is often used when writing SQL queries is FROM, SELECT, WHERE, ORDER BY, GROUP BY, and HAVING [12]. To make an effective query, students must first visualize where the data is and how to retrieve the data. Therefore, the SELECT syntax in SQL queries becomes the main concept that needs to be taught, especially when students have mastered the transition SELECT query to understand and explore other syntaxes more easily[13].

The topics most frequently discussed in the database teaching process are normalization, SQL syntax, operating database management tools and designing ERD [14]. Many studies have been carried out to improve students' understanding of database courses, including using a step-by-step approach, information systems with a role-playing approach [15], SQL teaching with problem based learning[16]. However, teaching with the existing approach is focusing on the theory and does not focus on students' ability to solve problems related to databases[14].

SQL query reconstruction has been implemented in a medical system that focuses on how to make a valid SQL query. The research discusses the concept of class, object, and data properties of a database [17]. Similar research created an E-Learning platform with the drag and drop concept. The study uses 30 participants concluded that the e-learning application provided a new experience in learning SQL. Using drag and drop concept has minimized errors due to syntax errors but is still open-ended where students can arrange query code blocks freely[18].

Another approach besides the open-ended approach is the closed-ended approach, this approach is widely used in the data collection process. Close-ended is widely used because it simplifies the process of filling out answers by participants[19]. The close-ended approach can be used for transferring the teacher's knowledge and understanding

to the student [20]. Examples of closed-ended questions are questions that already have some of the correct answers, then students fill in the rest of the questions that have not been filled in based on the available choices. This allows students to think logically and solve problems according to the context and scope of the questions given [21].

SQLearn is a new SQL query learning tool designed using combination between drag and drop and close-ended approach. In this tool student will learn to write SQL query by filling the gap using a drag and drop action to a predefined answer provided by the lecturer.

Our hypothesis for this study is by learning SQL query using SQLearn student will have a better skill in writing a SQL query. To measure student skill, we conduct an experiment with 28 participant and compare result between pre-test and post-test.

2 Literature Study

2.1 Close Ended

Close-ended questions is a method used to get response from respondent by limiting the answer to a set of offered alternatives [22]. By providing set of answer analyzing and processing close-ended questions is easier than open-ended questions. Close-ended questions also reduce the possibility of invalid answer[23]. Close-ended questions also enable respondent to respond with specific answer to the questions topic [24].

2.2 Drag And Drop

Drag and drop are widely used as user interface to learn programming language. Scratch [25] is a learning tool where a user can learn to create a program by dragging and dropping several graphical blocks into a scripting window. User can see the result of their blocks script interactively on the result area. Using graphical blocs help user focus on the learning process instead of typing complex syntax.

Several learning tools for database also using drag and drop style like Scratch, e.g., Bags, DB-Learn, DBSnap and ScrambleSQL. Bags using graphical blocks that provide data or relational operator for the user. User can drag and drop these blocks to a scripting area to create a database query. When the user executes the blocks, Bags will show result of the query in a table [26]. DB-Learn using similar concept with Bags with advanced cloud support [27]. DBSnap using a drag and drop blocks that create a tree instead of a linear blocks [28].DBSnap shows result of the block as a query, table result and node result. While Bags, DB-Learn and DBSnap focusing on relational algebra, ScrambleSQL focus on enhance learning for SQL command writing. ScrambleSQL provide list of SQL keyword, operator symbol and table as blocks that a student can drag to playground pane. Whenever student change the block in playground pane a SQL statement will be evaluated and show the result in the result table [18].

3 Sqllearn

SQLearn is a web-based SQL Query learning tool that applies close ended and drag and drop in the learning process. The lecturer is responsible to create a case study and questions. To create a case study lecturer must import a backup file that consist of SQL database schema and dataset for the case study. This database schema and dataset will be the base for checking query result created by student. The process of creating case study shown in Fig.1.

Tambah Studi Kasus

* Nama Studi Kasus

Studi Kasus . . .

Database

Click or drag file to this area to upload

Hanya bisa upload file .sql

Cancel Submit

Fig. 1. Create case study.

Edit Soal X

* Teks Soal

Dosen ingin menampilkan daftar kode jadwal teratur berdasarkan

* Kategori

Close-Ended v

Komponen SQL ?

SELECT jp_mulai jp_mulai jp kode_jp, X

+ Tambah komponen SQL

Petunjuk Jawaban ?

● Untuk bagian komponen yang kosong dapat diisi dengan ' _ ' (double-underscore)

_ _ _ FROM _ ORDER BY _ X

+ Tambah komponen petunjuk SQL

* Jawaban Benar

SELECT kode_jp, jp_mulai FROM jp ORDER BY jp_mulai

Fig. 2. Create questions and sql blocks.

When creating questions lecturer must set up several things the first one is the questions itself, lecturer must choose case study related to the questions, upload the image of desirable query result and setup key query answer for the questions. Key query answer consists of key and draggable SQL query parts blocks lecturer mark which parts of the key query answer that will be shown and which parts that will be set as draggable blocks. The process to create questions shown in Fig. 2.

Pertanyaan

Dosen ingin menampilkan daftar kode jadwal terurut berdasarkan jam mulai belajar

Daftar Tabel

▼ Tabel jadwal

Daftar Kolom :

- id-int(11)
- kode_jadwal-varchar(255)
- kode_kelas-varchar(255)
- kode_dosen-varchar(255)
- kode_mk-varchar(255)
- kode_ruang-varchar(255)
- kode_hari-varchar(255)
- jp_mulai-varchar(255)
- jp_selesai-varchar(255)

Preview Hasil Query

Soal Selanjutnya

Fig. 3. Question and table preview on student account.

SQL Query 16:16:43

Constructed SQL :
SELECT __ FROM __ ORDER BY __

Komponen SQL

jp_mulai	jp_mulai	jp	kode_jp,
----------	----------	----	----------

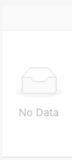
Jawaban SQL

SELECT		FROM		ORDER BY	
--------	--	------	--	----------	--

Test Query Reset Simpan Jawaban

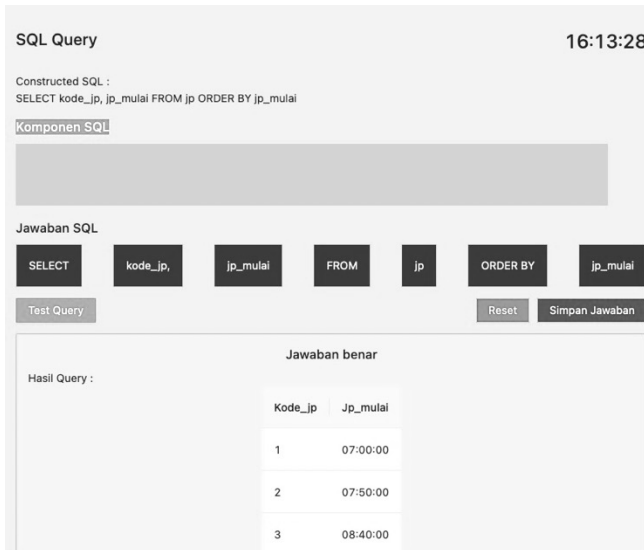
Hasil Query :

SyntaxError



No Data

Fig. 4. Error answer response.



The screenshot shows the SQL Query interface. At the top, it displays the time 16:13:28. Below that, the constructed SQL query is shown: `SELECT kode_jp, jp_mulai FROM jp ORDER BY jp_mulai`. A section titled "Komponen SQL" contains a list of draggable SQL blocks: `SELECT`, `kode_jp`, `jp_mulai`, `FROM`, `jp`, `ORDER BY`, and `jp_mulai`. Below the blocks are buttons for "Test Query", "Reset", and "Simpan Jawaban". The "Jawaban SQL" section shows the result of the query, which is a table with the following data:

Kode_jp	Jp_mulai
1	07:00:00
2	07:50:00
3	08:40:00

Fig. 5. Correct answer response.

When working on the questions student can see the questions, description of table that used for related questions, a preview of the expected query result, a pre constructed query answer and set of draggable SQL blocks. The student view on questions shown in Fig. 3 and Fig. 4.

Student can answer the question by dragging SQL blocks to the correct position on the preconstructed query answer. Before submitting the answer, student can test the query if the test query can be executed SQLearn will show the preview of query result if there are any error SQLearn will show a syntax error message. Students must arrange the SQL blocks to correct position in the pre constructed answer before continuing to the next questions. The student answer process shown in Fig. 5.

4 Experimental Settings And Discussion

4.1 Procedure

Twenty-eight sophomore students were involved in this experiment. They are from Information Technology department majoring Informatic Technic. They've learned the DBMS course subject. Makes them familiar with the basic of SQL query syntax. To handle the condition, then a complex sets of SQL syntax was created by the teacher to be used as the material in this experiment. The design of the experiment is shown in Fig.6.

During the experiment, the teacher was not explaining the material at all. The answer for both pre and posttest were not given to the students until all the cycles were finished. First activity, the students will learn the syntax from a textbook gave by the teacher. The material itself contains of explanation and example of the SQL query syntax. After reading, then a pre-test was conducted, during the test students were not allowed to open the

textbook. Next activity was using the application to help students understanding the syntax by them self. the last activity would be post-test. both pre and post-test were the same test.

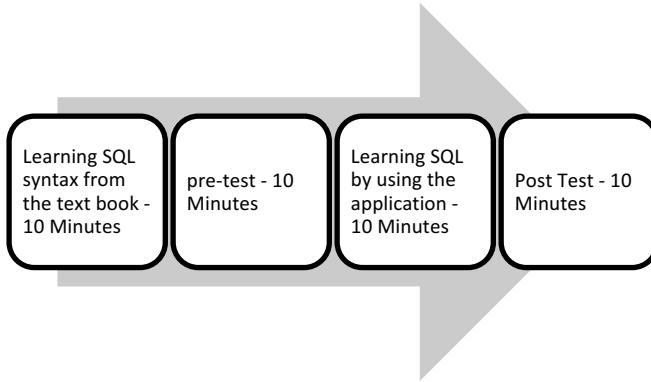


Fig. 6. Experiment design.

4.2 Result

Analyzing the experiment results was conducted to answer the research question. T-test statistical analysis was used to compare the test results because it could give a strong result to determine the differences between pre and post-test. Pearson correlation analysis was also used to measure the correlation between tests. Based on TABLE I, it shows that there are significant differences occur between the pre-test and post-test (p two tail = 2.01263E-08; $p < 0.001$), p -value were less than 0.001 means that the score gap between pre-test and post-test are significant different, where post-test is over-achieved the pre-test. Pearson correlation also show that a strong positive correlation appears ($0.5 < 0.721257512 < 1$) where 1 is the strongest.

4.3 Analysis

To support the result, box plot diagram was used to give a view on the density of score from both tests. Based on Fig. 7, it is clearly show that the post-test over-achieved the post test, the density of the score also wider and higher than the pre-test. Fig. 8, gave a better view by showing that most students are improved. Only few students are not improved, but they already had a good score.

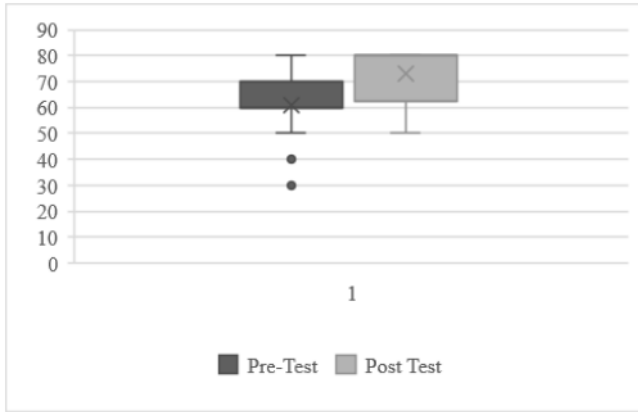


Fig. 7. Box plot for pre-test and post-test.

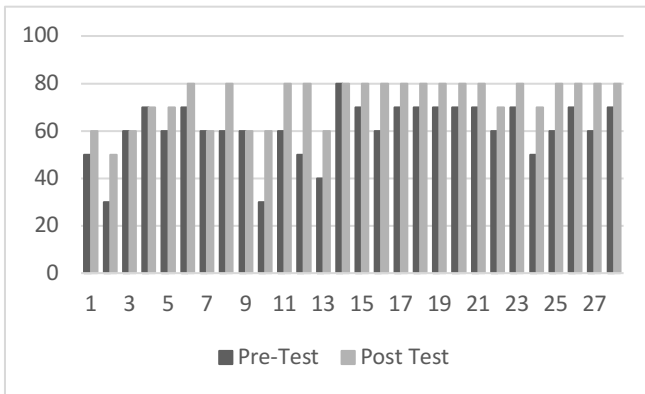


Fig. 8. Block chart for pre-test and post-test.

Table 1. Pearson correlation and t-test result.

t-Test:Paired Two Sample for Means	Pre-Test	Post Test
Mean	60.71428571	73.21429
Variance	147.6190476	89.28571
Observations	28	28
Pearson Correlation	0.721257512	
Hypothesized Mean Difference	0	
df	27	
t Stat	-7.833494518	
P(T<=t) one-tail	1.00632E-08	
t Critical one-tail	1.703288446	

P(T<=t) two-tail	2.01263E-08
t Critical two-tail	2.051830516

5 Conclusion

Based on the fact found in the experiments, there are strong indication showing that the used of the application could improve the post-test. most students, are able to improve their test scores. The used of the application had given them a better score means that their understanding are improved. One of the reasons of this improvement is because SQLearn provide them with some guidance so that they could analyzed their own understanding. The application also provide some suggestions that could be used to investigate the mistakes made by the students. Since the teacher didn't assist the students during the experiments then one of possible reason is the application could assist student for self-learning. The result of this experiment shows a positive result of the technology used in a learning environment. It gave a positive indication that the technology can be consider to be used in a real class situation.

References

1. M. Cvetanovic, Z. Radivojevic, V. Blagojevic, and M. Bojovic, "ADVICE—Educational System for Teaching Database Courses," *IEEE Trans. Educ.*, vol. 54, no. 3, pp. 398–409, Aug. 2011, doi: 10.1109/TE.2010.2063431.
2. D. Puspitasari, P. P. Arhandi, P. Y. Saputra, Y. W. Syaifudin, H. A. Himawan, and P. A. Sholihah, "Online judge MySQL for learning process of database practice course," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 523, p. 012046, Jul. 2019, doi: 10.1088/1757-899X/523/1/012046.
3. D. Nolan and D. Temple Lang, "Computing in the Statistics Curricula," *Am. Stat.*, vol. 64, no. 2, pp. 97–107, May 2010, doi: 10.1198/tast.2010.09132.
4. L. Hood, "Systems biology: integrating technology, biology, and computation," *Mech. Ageing Dev.*, vol. 124, no. 1, pp. 9–16, Jan. 2003, doi: 10.1016/S0047-6374(02)00164-1.
5. E. S. Adams and M. Granger, "Managing the Introductory Database Course: What Goes In and What Comes Out?," p. 2, 2004, doi: <https://doi.org/10.1145/1028174.971467>.
6. M. Murray, "Expanding the Database Curriculum," p. 8, 2008.
7. T. Taipalus, "A Notation for Planning SQL Queries," vol. 30, p. 9, 2019.
8. T. de Jong, "Cognitive load theory, educational research, and instructional design: some food for thought," *Instr. Sci.*, vol. 38, no. 2, pp. 105–134, Mar. 2010, doi: 10.1007/s11251-009-9110-0.
9. R. Kearns, S. Shead, and A. Fekete, "A teaching system for SQL," in *Proceedings of the second Australasian conference on Computer science education - ACSE '97*, The Univ. of Melbourne, Australia, 1996, pp. 224–231. doi: 10.1145/299359.299391.
10. A. Mitrovic, "A Knowledge-Based Teaching System for SQL," p. 6, 1998, doi: 10.1.1.31.1046.
11. J. Rith, P. S. Lehmayr, and K. Meyer-Wegener, "Speaking in tongues: SQL access to NoSQL systems," in *Proceedings of the 29th Annual ACM Symposium on Applied*

- Computing, Gyeongju Republic of Korea, Mar. 2014, pp. 855–857. doi: 10.1145/2554850.2555099.
12. A. Ahadi, J. Prior, V. Behbood, and R. Lister, “A Quantitative Study of the Relative Difficulty for Novices of Writing Seven Different Types of SQL Queries,” in Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education, Vilnius Lithuania, Jun. 2015, pp. 201–206. doi: 10.1145/2729094.2742620.
 13. P. Garner and J. Mariani, “Learning SQL in steps,” *J. Syst. Cybern. Inform.*, vol. 13, no. 4, pp. 19–24, 2015.
 14. K. S. Sastry, “An Effective Approach for Teaching Database Course,” p. 11, 2015.
 15. D. Shaw, B. J. Woodford, and G. L. Benwell, “Educating future IS professionals through real-world integration,” *Int. J. Teach. Case Stud.*, vol. 1, no. 1/2, p. 66, 2007, doi: 10.1504/IJTCS.2007.014210.
 16. A. S. Md. L. Hoque, G. Md. M. Bashiry, and Md. R. Uddin, “Equivalence of Problems in Problem Based e-Learning of Database,” in 2014 IEEE Sixth International Conference on Technology for Education, Amritapuri, Clappana P. O., India, Dec. 2014, pp. 106–109. doi: 10.1109/T4E.2014.23.
 17. H. Gorskis, “SQL query construction from database concepts,” in 2018 59th International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS), Riga, Oct. 2018, pp. 1–4. doi: 10.1109/ITMS.2018.8552953.
 18. C. Phewkum, J. Kaewchaiya, K. Kobayashi, and K. Atcharyachanvanich, “ScrambleSQL: A Novel Drag-and-drop SQL Learning Tool,” in 2019 23rd International Computer Science and Engineering Conference (ICSEC), Phuket, Thailand, Oct. 2019, pp. 340–344. doi: 10.1109/ICSEC47112.2019.8974815.
 19. V. Baburajan, J. de A. e Silva, and F. C. Pereira, “Open-Ended Versus Closed-Ended Responses: A Comparison Study Using Topic Modeling and Factor Analysis,” *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 4, pp. 2123–2132, Apr. 2021, doi: 10.1109/TITS.2020.3040904.
 20. B. S. Andoko, Y. Hayashi, T. Hirashima, and A. N. Asri, “Improving English reading for EFL readers with reviewing kit-build concept map,” *Res. Pract. Technol. Enhanc. Learn.*, vol. 15, no. 1, p. 7, Dec. 2020, doi: 10.1186/s41039-020-00126-8.
 21. W.-L. Lin and Y.-W. Lien, “The Different Role of Working Memory in Open-Ended Versus Closed-Ended Creative Problem Solving: A Dual-Process Theory Account,” *Creat. Res. J.*, vol. 25, no. 1, pp. 85–96, Jan. 2013, doi: 10.1080/10400419.2013.752249.
 22. W. Foddy, *Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research*, 1st ed. Cambridge University Press, 1993. doi: 10.1017/CBO9780511518201.
 23. U. Reja, K. L. Manfreda, V. Hlebec, and V. Vehovar, “Open-ended vs. Close-ended Questions in Web Questionnaires,” p. 19.
 24. A.-H. Chen, N.-F. A. Bakar, and C. S.-Y. Lam, “Comparison of open-ended and close-ended questions to determine signs and symptoms of eye problems among children,” *J. Optom.*, vol. 13, no. 2, pp. 81–87, Apr. 2020, doi: 10.1016/j.optom.2019.07.002.
 25. M. Resnick et al., “Scratch: Programming for Everyone,” p. 12.
 26. J. Gorman, S. Gsell, and C. Mayfield, “Learning relational algebra by snapping blocks,” in Proceedings of the 45th ACM technical symposium on Computer science education, Atlanta Georgia USA, Mar. 2014, pp. 73–78. doi: 10.1145/2538862.2538961.
 27. R. Vinayakumar, K. Soman, and P. Menon, “DB-Learn: Studying Relational Algebra Concepts by Snapping Blocks,” in 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Bangalore, Jul. 2018, pp. 1–6. doi: 10.1109/ICCCNT.2018.8494181.

28. Y. N. Silva and J. Chon, “DBSnap: Learning Database Queries by Snapping Blocks,” in Proceedings of the 46th ACM Technical Symposium on Computer Science Education, Kansas City Missouri USA, Feb. 2015, pp. 179–184. doi: 10.1145/2676723.2677220.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

