ARITHMETICNORM FOR LEARNING DATABASE NORMALIZATION



UNIVERSITI TEKNIKAL MALAYSIA MELAKA

BORANG PENGESAHAN STATUS LAPORAN

JUDUL: ARITHMETICNORM FOR LEARNII	NG DATABASE NORMALIZATION
SESI PENGAJIAN: Sesi 2020 / 2021	
Saya: <u>AISHA HASAN SAIF AL-BAKRI</u>	
mengaku membenarkan tesis Projek Sarjana M Teknikal Malaysia Melaka dengan syarat-syarat	<u> </u>
 Tesis dan projek adalah hakmilik Universe. Perpustakaan Fakulti Teknologi Maklu salinan unituk tujuan pengajian sahaja. Perpustakaan Fakulti Teknologi Maklu salinan tesis ini sebagai bahan pertukaran. * Sila tandakan (✓) TERHAD	mat dan Komunikasi dibenarkan membuat mat dan Komunikasi dibenarkan membuat
dsha	
(TANDATANGAN PELAJAR)	(TANDATANGAN PENYELIA)
Alamat tetap: <u>28, Jalan Merdeka 13</u>	
Taman Merdeka, Batu Berendam	Nama Penyelia
Tarikh:12/9/2021	Tarikh:
CATATAN: * Jika tesis ini SULIT atau TEl berkuasa.	RHAD, sila lampirkan surat daripada pihak

ARITHMETICNORM FOR LEARNING DATABASE NORMALIZATION

AISHA HASAN SAIF AL-BAKRI



This report is submitted in partial fulfillment of the requirements for the Bachelor of Computer Science (Database Management) with Honours.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY UNIVERSITI TEKNIKAL MALAYSIA MELAKA

DECLARATION

I hereby declare that this project report entitled

ARITHMETICNORM FOR LEARNING DATABASE NORMALIZATION

is written by me and is my own effort and that no part has been plagiarized without citations.

WALAYSIA OSha		
STUDENT :	Date : _	12 /9 /2021
AISHA HASAN SAIF AL-BAKRI	1	
برسيتي تيكنيكل مليسياً ملاك	اونيؤ	
LINIVERSITI TEKNIKAL MALAYSIA MEL	ΔΚΔ	

I hereby declare that I have read this project report and found this project report is sufficient in term of the scope and quality for the award of Bachelor of Computer Science (Database Management) with Honours.

	مو		
SUPERVISOR	:	Date : _	12 /9 /2021
	TS.HIDAYAH BINTI RAHMALAN		

DEDICATION

I dedicate this project to my beloved parents who's mean so much to me and helped me in all things great and small.

I also dedicate this work to my respected supervisor, Ts. Dr. Hidayah binti Rahmalan who encouraged me and pushed me to complete the project.

All of you have been my best cheerleaders. Without them after Allah, I would not have achieved this success.

I will always appreciate all they have done; may Allah grant you all Jannah Firdaus.

ACKNOWLEDGEMENTS

In the name of Allah, the most gracious and the most merciful. Alhamdulillah, all praises to Allah for the strengths and his blessing in completing my final project.

First and foremost, I express my heartfelt gratitude to my beloved father and my beloved mother for their unwavering love and support throughout the years. Thank you for guiding me in learning and doing everything in my power to help me walk along the path of greatness. You sacrificed many things for me to provide me with all the happiness. Whatever success I have now is due to both of you. Both of you aided me emotionally and monetarily.

Special appreciation goes to my respected supervisor Ts. Dr. Hidayah Binti Rahmalan for assisting me in developing my technical abilities and supporting me throughout my project with her patience and knowledge while enabling me to work in my own way. Her support and persistence are responsible for my bachelor's degree, and without her, this project would not have been accomplished. There is not a better or kinder supervisor in the world.

Special thanks to my wonderful family for being a part of my life and being another huge part in my study and my rock when I needed it the most and for their supporting for every decision I make.

In my daily study, I have been blessed with a friendly group of friends and lectures. Thanks to all my lecturers and friends for their cooperation and their invaluable assistance.

ABSTRACT

Normalization, which is a systematic way of decomposing tables, is the most extensively used process for examining relational databases. Its purpose is to design a set of relational tables with the least amount of data redundancy possible while retaining consistency. It also prevents undesirable features like insertion, update, and deletion anomalies. Thus, it is important for students to correctly implement the database normalization process. However, some students may have trouble with database normalization, which can lead to data duplication and the creation of databases that are larger than they need to be, slowing access. From our preliminary survey, 60% among FTMK lecturers agree that normalization is the most difficult topic in database and 90% agree that students have difficulties in learning normalization topis. While 81% among students agree that normalization is a difficult topic and 22% agree that normalization is the most difficult topic in database. Due to this matter, our study proposed a normalization tool named as ArithmeticNorm. Our ArithmeticNorm tool provides an environment for learning normalization of relational database to give students an interactive hands-on experience in database normalization process. From a total of 32 ArithmeticNorm users, 93% found that our tool will increase their productivity and 84% agree that ArithmeticNorm system will be useful in their studies.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

ABSTRAK

Normalisasi, yang merupakan cara yang sistematik untuk menguraikan jadual, adalah proses yang paling banyak digunakan untuk memeriksa pangkalan data relasi. Tujuannya adalah untuk merancang satu set jadual hubungan dengan sekurang-kurangnya jumlah data yang mungkin dan kemungkinan mengekalkan konsistensi. Ia juga menghalang ciri yang tidak diingini seperti penyisipan, kemas kini, dan penghapusan anomali. Oleh itu, adalah penting bagi pelajar untuk melaksanakan proses normalisasi pangkalan data dengan betul. bagaimanapun, sesetengah pelajar mungkin menghadapi masalah dengan normalisasi pangkalan data, yang boleh menyebabkan duplikasi data dan penciptaan pangkalan data yang lebih besar daripada yang mereka perlukan, melambatkan akses. Daripada tinjauan awal kami, 60% daripada pensyarah FTMK bersetuju bahawa normalisasi adalah topik yang paling sukar dalam pangkalan data dan 90% bersetuju bahawa pelajar mengalami kesukaran dalam mempelajari topis normalisasi. Walaupun 81% di kalangan pelajar bersetuju bahawa normalisasi adalah topik yang sukar dan 22% bersetuju bahawa normalisasi adalah topik yang paling sukar dalam pangkalan data. Oleh kerana perkara ini, kajian kami mencadangkan alat normalisasi yang dinamakan sebagai ArithmeticNorm. Alat aritmetik kami menyediakan persekitaran untuk pembelajaran normalisasi pangkalan data relasi untuk memberi pelajar pengalaman yang interaktif dalam proses dalam proses normalisasi pangkalan data. Daripada sejumlah 32 pengguna aritmetik, 93% mendapati bahawa alat kami akan meningkatkan produktiviti mereka dan 84% bersetuju bahawa sistem aritmeticnorm akan berguna dalam pengajian mereka.

Table of Contents

SUBJECTPA	1GE
DECLARATION	II
DEDICATION	III
ACKNOWLEDGEMENT	IV
ABSTRACT	V
ABSTRAK	VI
TABLE OF CONTENTS	
LIST OF TABLES	
LIST OF FIGURES	XIII
MALAYSIA	
CHAPTER I INTRODUCTION 1.1 Project Background	
1.2 Problem Statement	3
1.3 Objective	5
1 4 Project Scope	5
1.4 Project Scope	۔
1.4.2 Modules of System MALAYSIA MELAKA	6
1.5 Project Significant	7
1.6 Expected Output	7
1.7 Conclusion	8
CHAPTER II PROJECT METHODOLOGY & PLANNING	
	0
2.1 Introduction	
2.2 Project Methodology	10
2.2.1 Development Methodology	10
2.2.2 Database Methodology	11

2.2.3 Requirements Analysis Phase	12
2.2.4 Logical Design Phase	12
2.2.5 Physical Design Phase	12
2.2.6 Implementing Phase	13
2.2.7 Testing and Evaluation Phase	13
2.2.8 Maintenance Phase	13
2.3 Literature Review	14
2.4 Preliminary Survey	20
2.5 Project Schedule and Milestones	32
2.6 Conclusion	32
CHAPTER III ANALYSIS 3.1 Introduction 3.2 Problem Analysis 3.3 Proposed Improvement 3.3.1 Upload Table	34
3.3.2 Find Primary Key MIKAL MALAYSIA MELAKA	
3.3.3 Find Composite Key	41
3.4 Requirement Analysis of the To-Be System	43
3.4.1 Function Requirement	43
3.4.2 Non-Function Requirement	44
3.4.3 Other Requirement	45
3.4.3.1 Hardware Requirement	46
3.4.3.2 Software Requirement	46
3.5 Research Methodology	47
3.6 Conclusion	49

CHATER IV DEGIGN	
4.1 Introduction	50
4.2 System Architecture Design	51
4.3 Database Design	52
4.3.1 Conceptual Design	52
4.3.2 Logical Design	53
4.3.2.1 Data Dictionary	54
4.3.2.2 Database Normalization	57
4.3.3 Physical Design	60
4.4 Graphical User Interface Design (GUI)	61
4.5 Conclusion	73
CHATER V IMPLEMENTATION 5.1 Introduction	
5.2 System Development Environment Setup	75
5.2.1 Software Environment Setup	75
5.2.2 Database Environment Setup	78
5.2.2 Database Environment Setup 5.3 Database Implementation 5.4 Conclusion	79
5.4 Conclusion	80
CHATER VI TESTING	
6.1 Introduction	87
6.2 Test Plan	88
6.2.1 Test Organization	88
6.2.2 Test Environment	89
6.2.3 Test Schedule	89
6.3 Test Straregy	90
6.3.1 Classes of Test	92
6 / Test Design	03

6.4.1 Test Description
6.4.2 Test Data
6.5 Test Result and Analysis
6.6 Conclusion
CHATER VII CONCLUSION
7.1 Introduction
7.2 Observations on Weaknesses and Strengths
7.2.1 Weaknesses
7.2.2 Strengths
7.3 Propositions for Improvements
7.4 Contributions
7.5 Conclusion
REFERENCES 111
APPENDIX A: Normal Form Algorithm
APPENDIX B: Ouestions of Students' Questionnaire
APPENDIX C: Questions of Teachers' Questionnaire
APPENDIX D: Questions of Users' Questionnaire

LIST OF TABLES

Table 2. 1: Summary of Literature Review	18
Table 2. 2: Weighted Averages of the Scale Experience of Database Topic	s 28
Table 2. 3: project Milestones	31
Table 4. 1: Data Dictionary for User Table	54
Table 4. 2: Data Dictionary for Exam_paper Table	54
Table 4. 3: Data Dictionary for Results Table	55
Table 4. 4: Data Dictionary for Question Table	55
Table 4. 5: Data Dictionary for Choice Table	
Table 4. 6: Data Dictionary for Theory Table	56
Table 4. 7: Data Dictionary for Images Table Table 4. 8: Data Dictionary for Videos Table	56
Table 4. 8: Data Dictionary for Videos Table	56
Table 4. 9: Data Dictionary for Csvfile Table	57
Table 4. 10: Data Dictionary for Audio Table Table 4. 11: Stored Procedure	57
Table 4. 12: Stored Trigger KNIKAL MALAYSIA MELAKA	61
Table 6. 1: Test Environment Setup	89
Table 6. 2: Test Schedule for ArithmeticNorm System	90
Table 6. 3: Testing Approaches	91
Table 6. 4: Login Module Testing	94
Table 6. 5: Add Record Module	94
Table 6. 6: Add Media Data Module	95
Table 6. 7: Score Student Marks Module	95
Table 6. 8: Normalize Student Table Module	96
Table 6. 9: Test Data for Login Module	96
Table 6. 10: Test Data for Add Record Module	97
Table 6 11: Test Data for Add Media Data Module	90

Table 6. 12: Test Data for Score Student Marks Module	98
Table 6. 13: Test Data for Normalize Student Table Module	99
Table 6. 14: Test Result and Analysis for Login Module	99
Table 6. 15: Test Result and Analysis for Add Record Module	100
Table 6. 16: Test Result and Analysis for Add Media Data Module	100
Table 6. 17: Test Result and Analysis for Score Student Marks Module	101
Table 6. 18: Test Result and Analysis for Normalize Student Table Module	101
Table 6. 19: Weighted Averages of Usability and User Satisfaction	104



LIST OF FIGURES

	IAGI
Figure 2. 1: Rapid Application Development	10
Figure 2. 2: Database Life Cycle (DBLC)	11
Figure 2. 3: Normalization Tool by Griffith University [30]	19
Figure 2. 4: Gender of the Respondents	
Figure 2. 5: Range of Age of the Respondents	21
Figure 2. 6: Latest education level of the Respondents	21
Figure 2. 7: Semester that Respondents Taken Database Subject	21
Figure 2. 8: Difficulty of Normalization	22
Figure 2. 9: The Challenges Faced by The Students	22
Figure 2. 10: The Most Difficult Topic of Database Topics	23
Figure 2. 11: Agreement of Having a Normalization Tool	23
Figure 2. 12: Agreement of the Benefits of a Normalization Tool	
Figure 2. 13: Gender of the Teacher's Respondents	
Figure 2. 14: Range of Age of the Teacher's Respondents	25
Figure 2. 14: Range of Age of the Teacher's Respondents	26
Figure 2. 16: Duration of Teaching Database Subject	
Figure 2. 17: Teaching Normalization Topic	27
Figure 2. 18: Agreement on Facing Difficulty in Learning Normalization	n by
Students	27
Figure 2. 19: Scale of Experience	29
Figure 2. 20: The Most difficult Topic of Database Topic for the Teachers	29
Figure 2. 21: Experience using Normalization Tool	30
Figure 2. 22: Agreement of the benefit of a Normalization Tool	30
Figure 2. 23: Gantt Chart of ArithmeticNorm system	32
Figure 3. 1: Current System Flowchart	35
Figure 3. 2: Proposed Improvements Flowchart	36
Figure 3 3: Unload Table Flowchart	38

Figure 3. 4: Find Primary Key	40
Figure 3. 5: Find Composite Key Flowchart	42
Figure 3. 6: Research Methodology Phases	47
Figure 4. 1: Three-Tier Architecture	51
Figure 4. 2: Entity Relation Diagram (ERD) For ArithmeticNorm System	53
Figure 4. 3: ArithmeticNorm System Home Page	62
Figure 4. 4: ArithmeticNorm System Home Page (About)	62
Figure 4. 5: ArithmeticNorm System Home Page (Contact Us)	63
Figure 4. 6: Sign In page	63
Figure 4. 7: Registration page	64
Figure 4. 8: Theory Page	64
Figure 4. 9: Practice Page	65
Figure 4. 10: Student Home Page	
Figure 4. 11: User Profile Page	66
Figure 4. 12: Admin Home Page	66
Figure 4. 13: Upload Video Page	
Figure 4. 14: Add Theory Page	67
Figure 4. 15: Update Profile Page	68
Figure 4. 16: Upload Audio Page	68
Figure 4. 17: Test Yourself Page	69
Figure 4. 18: Results Page	69
Figure 4. 19: Dashboard Page	70
Figure 4. 20: Analysis of Users	70
Figure 4. 21: Analysis of Tables	71
Figure 4. 22: Analysis of Results	71
Figure 4. 23: Analysis of Exam Papers	72
Figure 4. 24: Analysis of Media Data	72
Figure 5. 1: XAMPP Installation Start	75
Figure 5. 2: Select Components of Server	76
Figure 5. 3: Select the Installation Folder Location	76
Figure 5. 4: Installation Process	77
Figure 5. 5: Completing the XAMPP setup	77

Figure 5. 6: XAMPP Control Panel	8
Figure 5. 7: PhpMyAdmin Interface	9
Figure 5. 8: Create Table Clause	80
Figure 5. 9: Select Statement for user Table	80
Figure 5. 10: SELECT Statement with WHERE Clause8	1
Figure 5. 11: Trigger After Insert to Question Table	31
Figure 5. 12: Trigger After Update on Question Table	32
Figure 5. 13: Trigger Before Delete from Question Table	32
Figure 5. 14: Trigger After Insert to csvfile Table	3
Figure 5. 15: Stored Procedure to Insert Results Data	3
Figure 5. 16: Stored Procedure to Insert User Data	4
Figure 5. 17: Stored Procedure to Insert Theory Data8	4
Figure 5. 18: Stored Procedure to Insert Video Data	5
Figure 5. 19: Stored Procedure to Insert CSV File	
Figure 5. 20: Stored Procedure to Insert Question Data8	6
Figure 6. 1: Gender of the Respondents User's	2
Figure 6. 2: Range Age of the Respondents User's	2
Figure 6. 3: Education Level of the Respondents User's	
Figure 6. 4: Learning/Using Database of the Respondents User's10	
Figure 6. 5: Benefits of ArithmeticNorm system10	15

CHAPTER I



Data is essentially the plain facts and statistics collected during the operations of a business. They can be used to measure/record a wide range of business activities - both internal and external. While the data itself may not be very informative, it is the basis for all reporting and as such is crucial in business.

Data normalization removes a variety of irregularities that might make data analysis more difficult. Anomalies might arise because of losing data, adding new information, or changing current data. Different benefits can be realized through other uses of the data and data analytics once the errors have been identified and eliminated from the system. Through data normalization, the information in a database can be formatted in such a way that it can be visualized

and analyzed.¹ Without it, a business can collect all the data it wants, but most of it would go unused, taking up space and providing little value to the organization. As a result, while attempting to load an integrated conceptual model into a database management system (DBMS), several issues may arise. These inconsistencies are caused by relationships that are created straight from user views.

There are several studies show the importance of database normalization and confirm the benefits of normalization for database. Mendjoge (2016) highlighted that it is critical for students to correctly implement the database normalization process. Codd (1970) confirm that the value of database normalization has been demonstrated and recommended in the literature for decades. While Ringlead (2019) highlighted that normalize your database before take action has numerous advantages for your sales. A database must be normalized to reduce redundancy (duplicate data) and ensure that only related data is stored in each table Li (2019). Winters (2021) said "normalization is important to protect data, correct data, and clean the database". While Muñoz (2019) said "It is more crucial to organize your company or catalog information than it is to find that pair of socks in your drawer every morning". Yee (2016) highlighted that the benefits of employing normalization are substantial, while the disadvantages are unavoidable in some situations. Normalization is important because it is extremely difficult to store items in a relational database that keeps the same information in many places Wikipedia (2018).

Therefore, various methods for improving teaching and learning database normalization were designed and developed. ArithmeticNorm system gives students an interactive hands-on experience in database normalization process starts from 1NF, 2NF, and finally 3NF by using certain techniques to determine a set of keys that exist in the relational database schema and proceed with the process of normalization. The goal of this study was to overcome the lack of knowledge and the teaching strategy used to improve learning database normalization.

¹ Import.io. (May 2019).

1.2 Problem Statement

This research starts with identifying problems and issues related to normalization. In this section, the problem statement was based on research materials such as journals or academic papers, and also preliminary findings through our survey.

There has been a lot of research done to show the complexity in the database normalization process, and how this is a tough topic not just for new database designers and developers, but also for university students studying this process. In this regard, Shanardi (2018) confirmed that normalizing a database is a complex and tough topic since analysts must understand the database's purpose. Brumm (2017) highlighted that database normalization concept can be hard to understand. Narvekar et al. (2016) observed that while taking a database management course, students find the processes of normalization and the ideas of normal forms is difficult to. Cooper (2021) said, "I found normalization quite difficult early in my job, likely because of my lack of experience".

According to Georgiev, N. (2008), computer science students have difficulty learning the fundamentals or concepts of database normalization, and difficulties in motivating them to learn database normalization because the subject appears to students to be dry and tough. Wang et al. (2010) highlighted that, because of the dry and academic manner in which database normalization is presented in textbooks and classes, it can be tough to motivate students to learn normalization topics.

Meanwhile, from our finding, normalization tends to be complex for new students and designer which could lead to difficulties that students face in learning database normalization. Thus, there are several problems that occur such as: i. Lose interest and trust in programming.

Due to that student may get poor database design, students will lose interest and trust in their programming and system they develop. Thus, that could lead students to do their work with less confidence.

ii. Current normalization tools do not provide a clear understanding of normalization.

Current algorithms that are used in available database normalization tools do not provide a clear understanding of normalization.

iii. Current normalization tools do not provide a visual aid for normalization.

The functional dependencies (FDs) that user input in available database normalization does not has data that show the user why his table needs to be separated into a set of tables. As a result, these tools do not provide a visual aid for normalization process.



According to the teacher's and student's preliminary survey results that can be referred in Section 2.4, normalization was confirmed the toughest and the most difficult topic in the database subject for students and for teachers as well. From a total of 10 respected teachers who teach database subjects, majority of them have chosen that normalization is a difficult and tough topic. While from a total of 97 students who took database subject, 79 students have chosen that normalization is a difficult topic. This becomes our interest to help students in learning normalization.

1.3 Objective

ArithmeticNorm system aims to overcome students' difficulties that they face while learning database normalization, the objectives of this project are as follows:

- i. To investigate students' and lecturers 'opinions about normalization topic.
- ii. To develop a tool that help student visualize normalization.
- iii. To evaluate the suggested normalization tool that has been developed.

1.4 Project Scope

The scope of this study is related to the topic of normalization. In this section, the scope also includes target users and module of system as shown in Section 1.4.1 and 1.4.2 respectively.

1.4.1 Target Users

The target user of ArithmeticNorm are computer science students, Undatabase designers, and those who are interested in the educational purposes of database normalization as well. Thus, the position of user that the system has is divided into two which are:

I. Student

Students can register to the system and start learning database normalization and practice their own database as well. so that users allow to insert and delete tables that are in CSV format, delete or update profile that include image media data, and view exams results.

II. Admin

The administrator manages all functions in the system and manages media data as well that could be added, updated, or deleted video and audio, do the functions that normal users have. In addition to getting an analysis of all users, tables, results, and media data.

1.4.2 Modules of System

There are 5 modules of ArithmeticNorm system which are:

I. Registration Module

This module will assist the student in registering and will make the manual registration process easier.

II. Login module

The authentication of user accounts will be aided by the login module. Users can access their accounts if they have a valid login username and password.

III. Assessment Module

This module will provide a testing environment where users can examine their knowledge gained in the database normalization.

IV. Theoretical Module

Theoretical module consists of theoretical explanations of normalization topics covering the normalization phase of the database design.

V. Practice module

This module will help students to practice database normalization through the system. ArithmeticNorm system guides students to perform a good database and normalize their table.

1.5 Project Significant

Due to the difficulties in learning normalization, this study intended to help and give a visual aid to the student in learning normalization. The suggested prototype that this study with to develop will benefit the students in their study and also in their career therefore, most of them would like to check the normalize of their database or to normalize their database as well.

According to the literature study table that will discover in Table 2.1, it has found that there are few commercial design tools are available to help students understand and learn normalization. These tools could be difficult for students to find out since it is not easily access. Therefore, ArithmeticNorm system expected to be more accessible, timely, and easy to understand by students in addition to solve student's problems that shown in Section 1.3.

1.6 Expected Output

ArithmeticNorm system expected to assess students' normalization knowledge by providing self-test questions. The system expected to provide media data in the education field to teach student normalization theories. ArithmeticNorm system also expected to accept students' CSV files, and push students to think about how to normalize their table by guide them to correct answers, and generate 1NF, 2NF, and 3NF as output. Students may have to practice or test his/her table.