Simulation and Modelling Project

A Project Documentation
Presented to Faculty of the Business Analytics Program of the
College of Business Administration
Silliman University
Dumaguete City, Philippines

In Partial Fulfillment of the Requirements of the Course
BA-AN 43P - Simulation Modeling and Analysis for Business

by

Renee Kristel Kate B. Faburada

January 2023

Table of Contents

I

DETERMINING THE TARGET GRADE USING GOAL SEEK

A. Introduction

Goal seeking is the process of finding the correct input value when only the output is known (Kenton, 2022). This is often done using the goal seek function in Microsoft Excel or what-if analysis via computer programs.

These can be great tools for students who are trying to figure out what grade they should get in a certain class in order to get a passing final grade. For college students in Silliman University, it is especially useful for those on certain scholarship programs where they are to maintain a certain total quality point average (TQPA), those vying for a spot in the honors or dean's list, or those who are just trying to pass.

This project aims to produce a web program that allows Silliman University college students to calculate their target grade for a specific subject based on the desired total quality point average (TQPA) using the principles of goal seek. The project uses a combination of PHP, SQL, and HTML to develop and implement a goal seek program that can be used in the web.

With this web program, the process of calculating target grades becomes automated. Students no longer have to go through the tedious process of manually calculating target grades and constantly adjusting their calculations to achieve realistic results.

B. Objective and Scope and Limitations

The main objective of this project is to produce a web program that allows Silliman University college students to calculate their target grade for a specific subject based on the desired total quality point average (TQPA) using the principles of goal seek.

This project is focused on the use of goal seek to calculate the target grade given a desired TQPA. The project is limited only to determining one target subject at a time. The web page is only applicable for Silliman University college students given that the main formula used in the calculation is based on the method used in calculating the university's college students' TQPA.

C. Review of Related Literature

Goal seeking is one of the tools used in "what-if analysis" on computer software programs. A what-if analysis is a process of changing values in (Microsoft Excel) cells to see how these changes will affect formula outcomes on the worksheet. In essence, one would be creating a scenario by asking "what if the output was X"—or basically, a cause and effect situation (Kenton, 2022).

In the case of this project, some possible question would be "What if the required TQPA was 3.25? What grade should I get in this class then?" or "What if the grade I need to graduate is 2.5?". GCF Global Learning (n.d.) and Excelbuddy.com (2019) clearly demonstrated how the Goal Seek function in Microsoft Excel to help students determine what grade they should strive for in a class requirement to achieve a certain grade.

However, Goal Seek works only with one variable input value (Microsoft, n.d.). Adding more variables, it becomes an optimization problem which uses other techniques such as linear programming.

This project uses PHP and SQL to develop a goal seek web program. Currently, there are no existing libraries or built in functions in PHP and SQL that can be utilized to automate the process.

D. Methodology

As the first step, a model for getting the target grade of a specific subject was first constructed using manual methods. PHP and SQL were used as the main programming languages for the foundation of this project. Since neither PHP nor SQL has a library for goal seek, an algorithm using the following formula was implemented to achieve a target grade:

$$Target\ Grade = ((TTQPA \times \sum U) / TU) - (\sum WG / TU)$$

where:

TTQPA = Target TQPA

$\sum U$ = Sum of All Units

TU = Number of Units of the Target Subject

$\sum WG$ = Sum of the Weighted Grade of All Subjects EXCEPT the Target Subject (Grade x Units)

All values are stored in a MySQL database table (Figure 1) with the following attributes:

subjectid: ID number automatically assigned to a subject

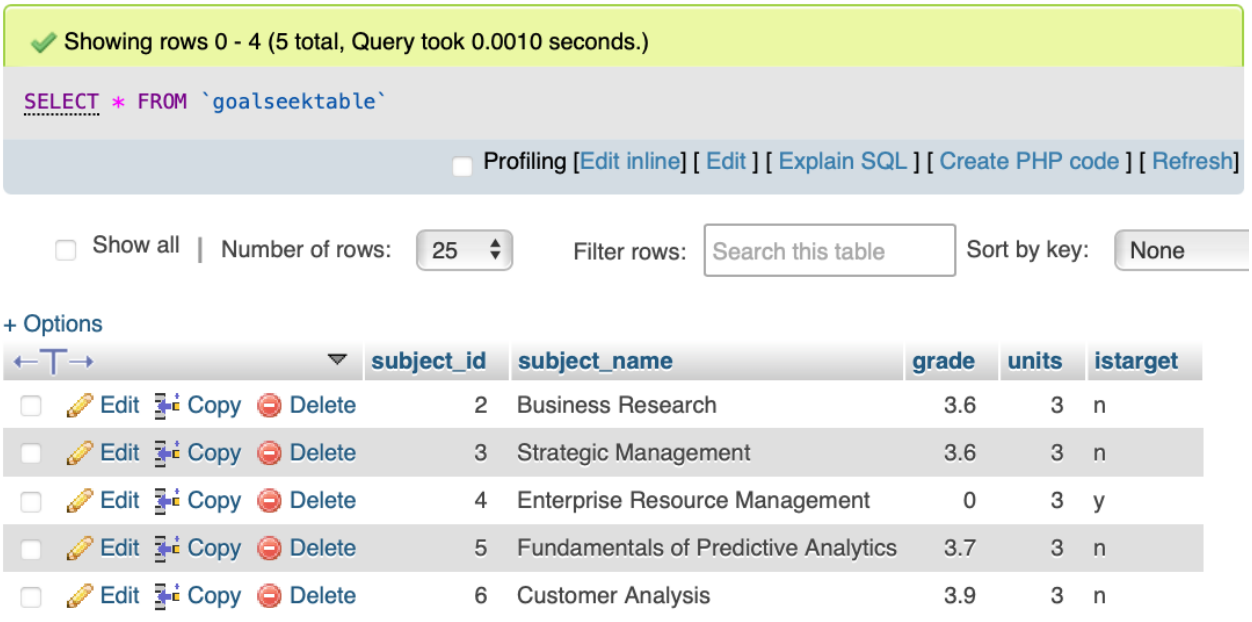
subjectname: Name of the subject

grade: Subject grade ranging from 0 to 4

units: The number of units assigned to a subject

istarget: This determines whether or not the subject is the target subject (y: the subject is the target; n: the subject is not the target).

The formula was implemented by creating a custom SQL query. PHP was used to send this query and other necessary queries to the MySQL database as well as retrieve results. Outcomes were compared with that of the results using manual calculations with the same values as a control to ensure accuracy.



	subject_id	subject_name	grade	units	istarget
Edit Copy Delete	2	Business Research	3.6	3	n
Edit Copy Delete	3	Strategic Management	3.6	3	n
Edit Copy Delete	4	Enterprise Resource Management	0	3	y
Edit Copy Delete	5	Fundamentals of Predictive Analytics	3.7	3	n
Edit Copy Delete	6	Customer Analysis	3.9	3	n

Figure 1. Grade Goal Seek Table in MySQL with Sample Values

With HTML input and select features, users are able to construct and modify their own grade table (Figure 2). Users should be able to add, delete, and edit subjects as well as reset the entire page to their liking. The users also have the option to switch target subjects and input their desired TQPA with a slider. The minimum and maximum range for the desired TQPA is based on the data inputted in the table. The initial TQPA serves as the lowest attainable TQPA. The highest attainable TQPA, on the other hand is based on the selected target subject and calculated in a way that the target subject grade does not exceed 4, which is the maximum grade for college students at Silliman University.

Appendix A explains in detail how the program runs.

E. Conclusion and Recommendation

Manually calculating grades and adjusting the equations to achieve a certain target can be a tedious process. This web program uses a combination of PHP, SQL, and HTML to automate the process of calculating target grades for Silliman University college students. It has features that allows students to create and modify their own grade tables and adjust target TQPA values and target subject grades just by using a slider. With the program, the students can also view the minimum and maximum TQPA they can achieve.

However, this only works for college level students in Silliman University. Future researchers can use this project as a guide in building other similar projects with the aim of automating the process of calculating the target grade.

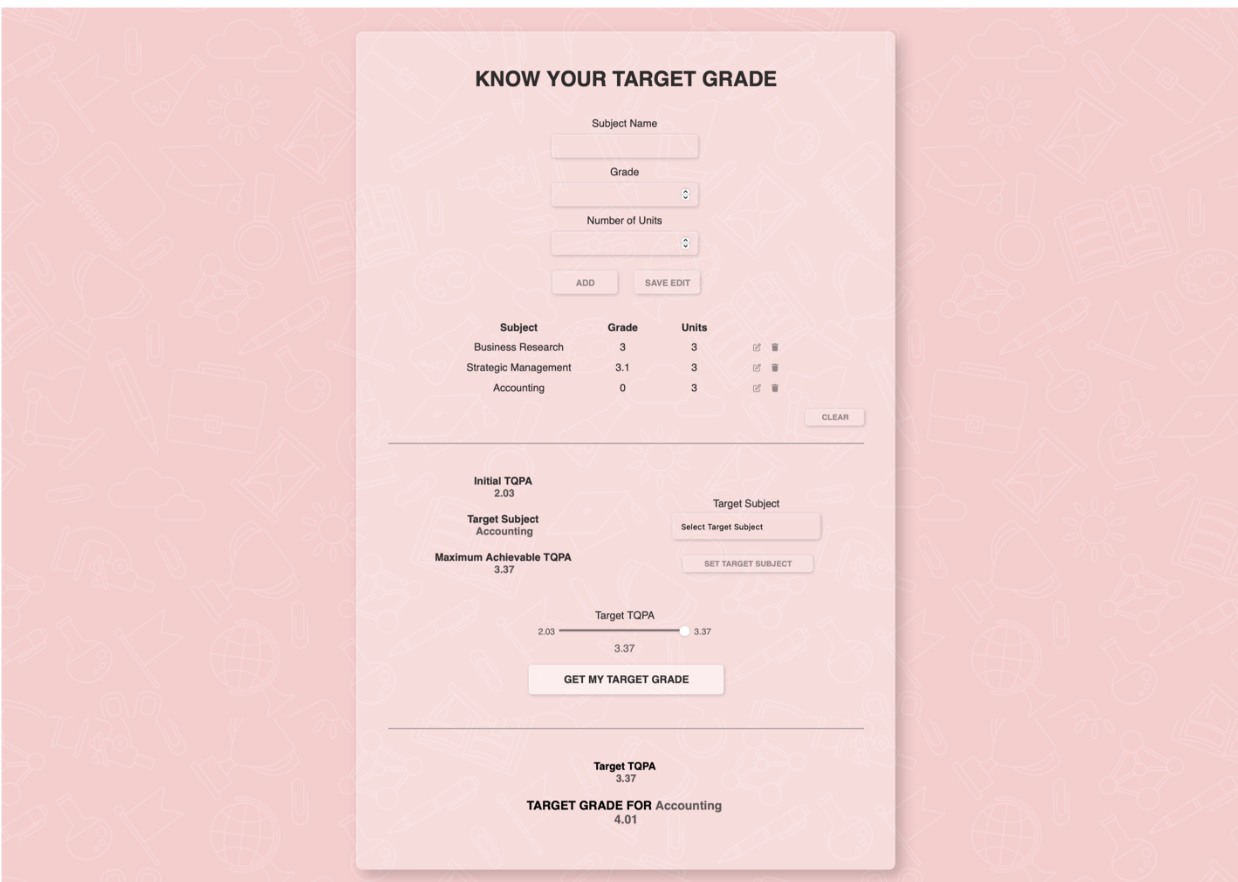


Figure 2. Grade Goal Seek Web View with Sample Values

II

MONTE CARLO SIMULATION OF GROSS PROFIT

A. Introduction

For any business, profit forecasting is crucial. Monte Carlo Simulation is used to provide the decision-maker with a range of potential outcomes and their probability of occurring for any course of action. It demonstrates the extreme scenarios—the results of taking a risk and taking the most cautious approach—along with all possible results for decisions that fall somewhere in the middle (Balogh et al., 2013). In financial forecasting, businesses should be able to get a better gauge of the likelihood that profits would fall within a certain range by using Monte Carlo Simulation (Tsoi, 2022).

Two of the most important indicators of health in a business is the cost of goods sold and the gross profit margin. The cost of goods sold is used to calculate the gross profit, where the gross profit is a useful ratio for businesses to see whether they are earning more than they spend on inventory. In the simplest sense, this shows the financial health of business operations (Rivero Gordimer & Company, 2021).

This project is aimed at developing a program for predicting the gross profit and possibility of achieving the target gross profit using Monte Carlo Simulation given that the average revenue and cost of goods sold, as well as their standard deviations, are known. JavaScript is the main language used to build the algorithm for running the simulations.

B. Objective and Scope and Limitations

The objective of this project is to develop a web program that predicts the gross profit and the possibility of achieving the targeted gross profit using Monte Carlo Simulation.

This project is focused on the use of Monte Carlo Simulation on predicting the gross profit and probability of achieving the target profit. The web program can be used by any merchandising business using different currencies. However, it can only be used under the assumption that the business already has access to historical data where the average revenue and average cost of goods sold, including their respective standard deviations, are known. Additionally, this project also assumes that the data is normally distributed.

C. Review of Related Literature

Monte Carlo Simulation is used to provide the decision-maker with a range of potential outcomes and their likelihood of occurring for any course of action. It demonstrates the most extreme scenarios—the results of taking a risk and taking the most cautious approach—along with all possible results for decisions that fall somewhere in the middle (Balogh et al., 2013).

Monte Carlo Simulation, also known as the Monte Carlo Method or a multiple probability simulation, is a mathematical technique, which is used to estimate the possible outcomes of an uncertain event (IBM Cloud Education, n.d.). Since its inception, Monte Carlo simulations have been used to assess the effects of risk in a variety of real-world contexts, including artificial intelligence, stock markets, sales and profit forecasting, project management, and pricing among others. Monte Carlo Simulations are usually used to model the probability of different outcomes that are not easily predictable due to the randomness in variables (Tsoi, 2022). In a Monte Carlo Simulation, a model is run at least a thousand times in order to obtain a series of results in which some are more likely to occur than others. They are then combined to form an overview from which conclusions and decisions are based upon.

This project focuses on the use of Monte Carlo Simulation in the prediction of gross profit and the probability of achieving a certain target given the revenue and cost of goods sold as the input values. The cost of goods sold and target profit margin according to Rivero Gordimer & Company (2021) are two of the most important indicators of health in a merchandising business. The cost of goods sold is used to calculate the gross profit, where the gross profit is a useful ratio for businesses to see whether they are earning more than they spend on inventory.

One of the assumptions in this project is normal distribution of data. Many industries and companies incorporate this type of distribution analysis into their business decision-making processes. Among the industries that also use this type of distribution in their analysis are sales and marketing, accounting, finance and stock market analysis, politics, logistics, and manufacturing (Indeed, 2022). Also known as the Gaussian distribution, the normal distribution is the most common type of distribution assumed in technical stock market analysis and in other types of statistical analyses. The standard normal distribution has two parameters: the mean and the standard deviation (Chen 2022).

D. Methodology

JavaScript was used as the main programming language to run the simulation. A model was first created using Microsoft Excel and then translated into code using JavaScript.

In the model, it is assumed that the two variables– the revenue and cost of goods sold– are random variables with a normal distribution and that their respective means and standard deviations are known. In Excel, values can be achieved using the following formula: =NORM.INV(probability, mean, standard deviation). With JavaScript, a function for a Gaussian distribution, which is normally distributed, was used. The function shared by Jóhannsson (2016) uses the principles of the Marsaglia polar method, a pseudo-random number sampling method for generating a pair of independent standard normal random variables.

Arrays of 1000 random variables for each the revenue and cost of goods sold was produced using a for loop and the following formula was applied on each pair of variables to determine the gross profit:

Gross Profit = Revenue – Cost of Goods Sold

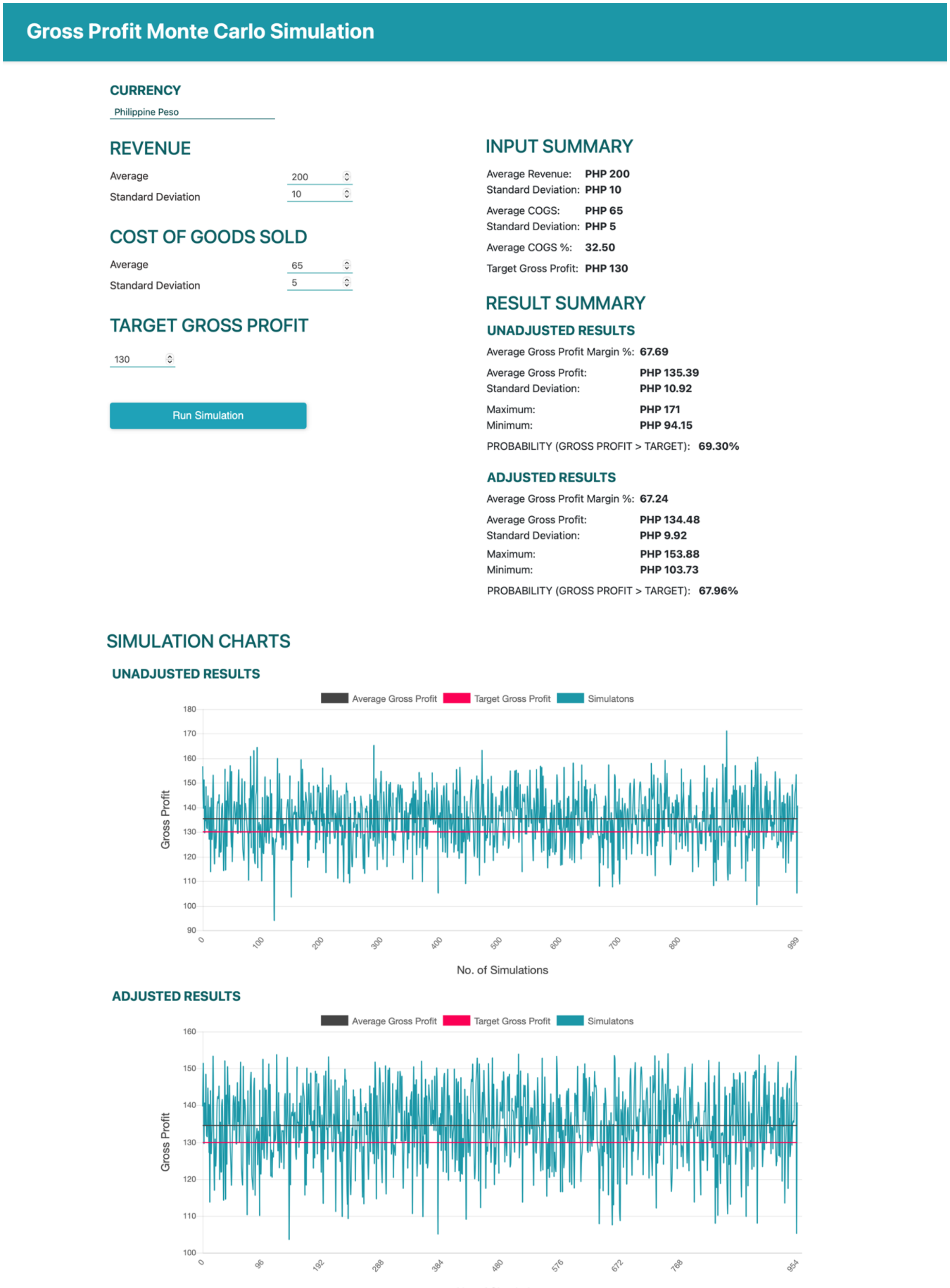


Figure 3. Gross Profit Monte Carlo Simulation Web View

In an interview with Skillings (2022), an Assistant Manager in Bellevue, Seattle's Home Depot, the top 10% and bottom 10% are considered as outliers which are values that have very little to no possibility of occurring. He suggested that in order to make the prediction more accurate the top and bottom 10% be removed.

With this, the array of 1000 values was adjusted to the suggested range, creating a new array and retaining the original array. The mean, standard deviation, maximum, minimum, and the probability of the predicted profit being above the target profit were then calculated.

HTML features that allow users to input values as well as review the information they have provided and view the result summary were added. For comparison purposes, both the adjusted and unadjusted results are displayed.

Two line graphs were constructed using Chart.js. The unadjusted graph features the all 1000 possible values for gross profit as well as the mean line and the target line. The adjusted graph features the values within the calculated range as well as the adjusted mean and target lines.

A more detailed explanation of the program is provided in *Appendix B*.

E. Conclusion and Recommendation

Monte Carlo Simulation is a useful technique in estimating the gross profit and probability of attaining a target goal profit. With this web program, merchandising companies or companies of similar nature should be able to get a better gauge of their company's financial health and make better informed decisions.

The model however, only offers a simplified and generalized view of the gross profit equation. This does not account for a detailed list of factors that go into the cost of goods sold. There are a lot of different variables that are put into the cost of goods sold and it varies from company to company. The researcher recommends a more in-depth exploration of these factors and consultation of professionals in the field in order to derive with a more accurate model.

III

PREDICTING A PATIENT'S LIKELIHOOD OF HAVING A HEART DISEASE USING A DECISION TREE

A. Introduction

Heart diseases, often used in exchange for cardiovascular diseases, are considered the leading cause of death. According to the World Health Organization (n.d.), an estimated 17.9 million people die from such diseases. Coronary heart disease, cerebrovascular disease, rheumatic heart disease, among other illnesses, are among the heart and blood vessel disorders collectively referred to as CVDs. Heart attacks and strokes account for more than four out of every five heart disease deaths, and they account for a third of the premature deaths among adults under the age of 70. Given such facts, it is important that such diseases are identified at an early stage to avoid further complications.

Multiple studies have been conducted on the use of machine learning in detecting heart diseases in patients (Chang et al., 2022; Nagavelli et al., 2022; Jindal et al., 2021; Shouman et al., 2011). This project uses a decision tree to determine whether or not a certain patient is likely to have any of the existing cardiovascular disease. While most studies use Python as the primary programming language, this project uses JavaScript to construct a decision tree given a widely used sample data.

For this project, 11 independent variables were considered: age, sex, type of chest pain experienced, resting systolic blood pressure, cholesterol, fasting blood sugar, resting ECG (electrocardiam) result, maximum heart rate, exercise-induced angina, ST depression induced by exercise, and ST slope at peak of exercise.

B. Objective and Scope and Limitations

The objective of this project is to be able to develop a web program that helps determine if a specific patient is likely to have any type of heart disease using a decision tree.

This project is focused on the use of a decision tree to predict an individual's likeliness of having a heart disease. The web program can be used by health professionals or patients who are already aware of their medical information based on previous check-ups. The only factors considered for predicting the likelihood of having a heart disease are the following 11 independent variables: age, sex, type of chest pain experienced, resting systolic blood pressure, cholesterol, fasting blood sugar, resting ECG (electrocardiam) result, maximum heart rate, exercise-induced angina, ST depression induced by exercise, and ST slope at peak of exercise.

C. Review of Related Literature

The World Health Organization (n.d) considers heart diseases as a leading cause of death worldwide. an estimated 17.9 million people die from such diseases. . Heart attacks and strokes account for more than four out of every five heart disease deaths, and they account for a third of the premature deaths among adults under the age of 70. It is important that illnesses such as heart diseases are identified at the earliest stage possible in order to avoid further complications which could be fatal. With today's technology, detecting diseases can now be done with the help of machine learning.

Machine learning is widely studied as a method of detecting heart diseases. Some of the techniques used by other researchers to diagnose heart diseases include neural networks, Naïve Bayes, genetic algorithms, decision trees, random forests, classification via clustering, direct kernel self-organizing map, and cross validation (Chang et al., 2022; Ghosh & Jana, 2022; Nagavelli et al., 2022; Jindal et al., 2021; Shouman et al., 2011)

This project applies the decision tree as the main tool for detecting heart diseases. Decision Trees, according to Scikit-learn.org, (n.d.) are a non-parametric supervised learning method used for classification and regression where the goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features.

With great success, researchers have been examining the use of decision trees in the identification of heart disease in patients. Andreeva (2006) in his study, used a decision in the diagnosing heart diseases with a 75.73% accuracy. Results in the study of Sitair-Taut et al. (2009) showed that when making an accurate prediction of heart disease, Naive Bayes and Decision Trees are not significantly different from one another after investigating both methods in the diagnosis. Additionally, in the study of Tu et al. it was concluded that in diagnosing heart diseases, using the decision tree showed 78.91% accuracy. However, the bagging algorithm showed better accuracy with a success rate of 81.41%. More recent studies involving the use of decision trees in predicting heart disease in individuals include that of Sai Krishna Redy et al. (2022), Maheswari & Pichay (2019), and Iliyas & Shaikh (2019). According to Sai Krishna Reddy et al. (2022) using a decision tree proved to be 82% accurate.

In the study of Ghosh and Jana (2022) on heart disease prediction using different classification models based on cross validation, the following variables were used: age, sex, chest pain type, resting blood pressure, cholesterol, fasting blood sugar, resting ECG results, maximum heart rate achieved, exercise-induced angina, old peak measured in ST depression, and the ST slope or the slope of peak exercise ST segment.

According to Finkelhor et al. (1986), "the ST segment shift relative to exercise-induced increments in heart rate, the ST/heart rate slope, has been proposed as a more accurate ECG criterion for diagnosing significant coronary artery disease." It is also stated by Istolahti et al. (2020) that "even minor ST depression in the electrocardiogram (ECG) is associated with cardiovascular disease and increased mortality."

D. Methodology

The dataset retrieved was created by combining different datasets already available independently but not combined before Fedesoriano (2021). The same data set was used in the project of Ghosh and Jana (2022).

There are originally 918 observations in the dataset described in *Figure 4.*

Attribute	Description	Type	Values
Age	age of the patient in years	Numerical	0-122 years
Sex	gender of the patient	Categorical	M - male F - female
ChestPainType	chest pain type	Categorical	TA - Typical Angina (chest pain caused by physical or emotional stress) ATA - Atypical Angina (does not have classical symptoms of chest pain) NAP - Non-Anginal Pain (chest pain not caused by a heart attack) ASY - Asymptomatic (no symptoms observed)
RestingBP	resting systolic blood pressure in mmHg	Numerical	50 to 370 mmHg

Cholesterol	serum cholesterol in mm/dl	Numerical	0 to 700 mm/dl
FastingBS	fasting blood sugar	Categorical	1 - if FastingBS > 120 mg/dl 0 - otherwise
RestingECG	resting electrocardiogram results	Categorical	Normal - normal ECG reading ST - ECG has ST wave abnormality LVH - showing probable or definite left ventricular hypertrophy by Estes' criteria
MaxHR	maximum heart rate achieved	Numerical	25 to 480 bpm
ExerciseAngina	exercise-induced angina	Categorical	Y - yes N - no
Oldpeak	ST depression of ECG	Numerical	
STSlope	the slope of the peak exercise ST segment	Categorical	Up - upward slope Flat - slope is zero Down - downward slope
HeartDisease	output class	Categorical	1 - individual has heart disease 2 - individual does not have heart disease

Figure 4. Heart Disease Dataset Information

The 11 independent variables in the dataset are: age, sex, type of chest pain experienced, resting blood pressure, cholesterol, fasting blood sugar, resting ECG (electrocardiam) result, maximum heart rate, exercise-induced angina, ST depression induced by exercise, and ST slope at peak of exercise. The target variable in the dataset to determine whether there exists a heart disease or not is represented by "HeartDisease".

Pre-processing was required to make the data usable. Although there were no missing values, there were values of 0 observed under cholesterol. It is a known fact that the Serum Cholesterol of a person can never be 0. Therefore, the rows given these values were dropped. We are now left with 746 rows.

JavaScript was used as the main programming language to construct a decision tree. The main framework for the decision tree algorithm is based on the project of Lahoduik (2018). The output of the framework was then compared to Python outputs using datasets from previous projects such as the recommended lenses dataset and results came out the same.

As the main file type of the dataset was in CSV format, it was converted to a JavaScript array using a converter then plugged into the main framework. HTML was used to display results on the web.

For ease of use, a form was created using HTML. Instead of having to use the fairly large decision tree, the user need only input the patient information and display the results on the same page showing the summary of the information given as well as whether or not the patient is likely to have a heart disease based on the inputted information. 1 indicates that the individual has a heart disease and 0 indicates that the individual does not have a heart disease. This was tested by inputting existing information from the dataset and comparing whether the results in the target variable "HeartDisease", is the same or not.

The web program uses CSS styling to create the tree design framework. *Appendix C* presents more details on how the program runs.

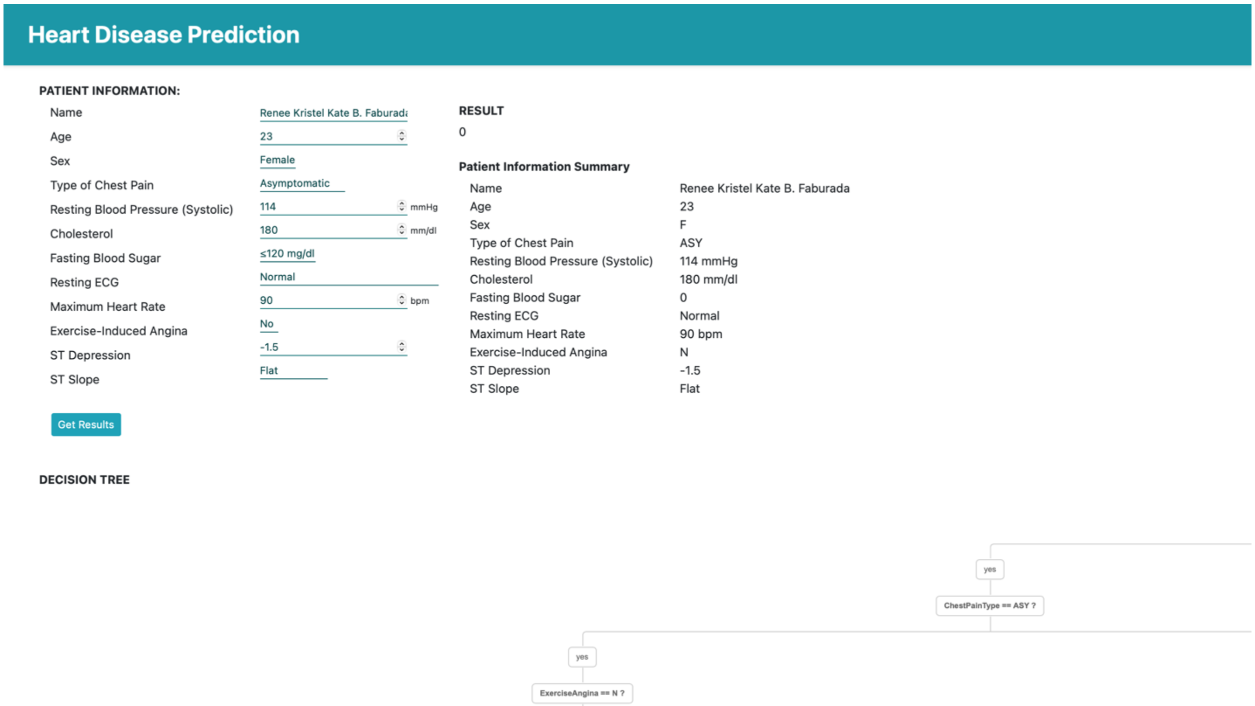


Figure 5. Heart Disease Prediction Web View

E. Conclusion and Recommendation

Heart disease is a serious condition that should be identified at an earliest stage possible. Based on multiple studies, machine learning is a helpful tool in detecting such diseases in patients. This project uses a decision tree to determine whether or not a certain patient or individual is likely to have a heart disease. However, this only covers the basics in constructing a decision tree prediction model.

Future researchers should further look into the accuracy of this model and consult experts when building their own.

IV
REFERENCES

Andreeva, P. (2006). Data Modelling and Specific Rule Generation via Data Mining Techniques. *International Conference on Computer Systems and Technologies (CompSysTech)*.

Balogh, P., Gorea, P., & Inceu, V. (2013). Profit Forecast Model Using Monte Carlo Simulation in Excel. *Romanian Statistical Review*.

Chang, V., Bhavani, V. R., Xu, A. Q., & Hossain, M. A. (2022). An artificial intelligence model for heart disease detection using machine learning algorithms. *Healthcare Analytics*, *2*, 100016. https://doi.org/10.1016/j.health.2022.100016

Chen, J. (2022, October 6). *Normal distribution: What it is, properties, uses, and Formula*. Investopedia. Retrieved December 9, 2022, from https://www.investopedia.com/terms/n/normaldistribution.asp

Excelbuddy.com. (2019, September 18). *Goal seek in excel – what-if analysis*. Excelbuddy.com. Retrieved December 8, 2022, from https://excelbuddy.com/goal-seek-in-excel-what-if-analysis/#school

Finkelhor, R. S., Newhouse, K. E., Vrobel, T. R., Miron, S. D., & Bahler, R. C. (1986). The St Segment/Heart Rate Slope as a predictor of coronary artery disease: Comparison with quantitative thallium imaging and conventional ST segment criteria. *American Heart Journal*, *112*(2), 296–304. https://doi.org/10.1016/0002-8703(86)90265-6

GCF Global Learning. (n.d.). *Excel: What-if analysis*. GCFGlobal.org. Retrieved December 8, 2022, from https://edu.gcfglobal.org/en/excel/whatif-analysis/1/

Ghosh, A., & Jana, S. (2022). A Study on Heart Disease Prediction using Different Classification Models based on Cross Validation Method. *INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT)*, *11*(6). https://doi.org/10.17577/IJERTV11IS060029

IBM Cloud Education. (n.d.). *What is Monte Carlo Simulation?* IBM. Retrieved November 4, 2022, from https://www.ibm.com/cloud/learn/monte-carlo-simulation

Iliyas, M. K., & Shaikh, I. S. (2019). Prediction of Heart Disease Using Decision Tree. *Allana Management Journal of Research, Pune*, *9*(2).

Indeed. (2022, March 29). *A guide to normal distribution and its uses in business*. Retrieved December 9, 2022, from https://ie.indeed.com/career-advice/career-development/normal-distribution

Istolahti, T., Nieminen, T., Huhtala, H., Lyytikäinen, L.-P., Kähönen, M., Lehtimäki, T., Eskola, M., Anttila, I., Jula, A., Rissanen, H., Nikus, K., & Hernesniemi, J. (2020). Long-term prognostic significance of the St Level and st slope in the 12-lead ECG in the general population. *Journal of Electrocardiology*, *58*, 176–183. https://doi.org/10.1016/j.jelectrocard.2019.12.010

Jindal, H., Agrawal, S., Khera, R., Jain, R., & Nagrath, P. (2021). Heart disease prediction using machine learning algorithms. *IOP Conference Series: Materials Science and Engineering*, *1022*(1), 012072. https://doi.org/10.1088/1757-899x/1022/1/012072

Jóhannsson, J. P. (2016). JavaScript Math.random Normal distribution (Gaussian bell curve)? Retrieved from https://stackoverflow.com/questions/25582882/javascript-math-random-normal-distribution-gaussian-bell-curve.

Kenton, W. (2022, October 28). *Goal Seeking*. Investopedia. Retrieved December 8, 2022, from https://www.investopedia.com/terms/g/goal-seeking.asp

Lahoduik, Y. (2018, January 11). *decision-tree-js*. Retrieved December 4, 2022 from https://github.com/lagodiuk/decision-tree-js.

Maheswari, S., & Pitchai, R. (2019). Heart disease prediction system using decision tree and naive Bayes algorithm. *Current Medical Imaging Formerly Current Medical Imaging Reviews*, *15*(8), 712–717. https://doi.org/10.2174/1573405614666180322141259

Microsoft. (n.d.). *Use Goal Seek to find the result you want by adjusting an input value*. Microsoft Support. Retrieved December 8, 2022, from https://support.microsoft.com/en-us/office/use-goal-seek-to-find-the-result-you-want-by-adjusting-an-input-value-320cb99e-f4a4-417f-b1c3-4f369d6e66c7

Nagavelli, U., Samanta, D., & Chakraborty, P. (2022). Machine learning technology-based heart disease detection models. *Journal of Healthcare Engineering*, *2022*, 1–9. https://doi.org/10.1155/2022/7351061

Platen, E., Schurz, H., & Kloeden, P. (1997). *Numerical Solution of SDE Through Computer Experiments*. Springer.

Rivero Gordimer & Company. (2021, May 7). *What is cost of goods sold (COGS) and why should you care?* Retrieved December 9, 2022, from https://www.rgcocpa.com/news/what-is-cost-of-goods-sold-cogs-and-why-should-you-care/

Sai Krishna Reddy, V., Meghana, P., Subba Reddy, N. V., & Ashwath Rao, B. (2022). Prediction on cardiovascular disease using decision tree and naïve Bayes classifiers. *Journal of Physics: Conference Series*, *2161*(1), 012015. https://doi.org/10.1088/1742-6596/2161/1/012015

scikit-learn.org. (n.d.). *1.10. decision trees*. scikit. Retrieved December 9, 2022, from https://scikit-learn.org/stable/modules/tree.html

Sitar-Taut, V. A. et al. (2009). Using machine learning algorithms in cardiovascular disease risk evaluation. *Journal of Applied Computer Science & Mathematics.*

Skillings, J. (2022, December 4). Gross Profit Montecarlo Simulation Review. personal.

Tsoi, T. (2022, January 26). *Application of monte carlo simulation in financial forecasting*. Medium. Retrieved November 6, 2022, from https://medium.com/@tsoiyingkit/application-of-monte-carlo-simulation-in-financial-forecasting-1ddb231080f9

Tu, M. C. et al. (2009). "Effective Diagnosis of Heart Disease through Bagging Approach." Biomedical Engineering and Informatics, IEEE.

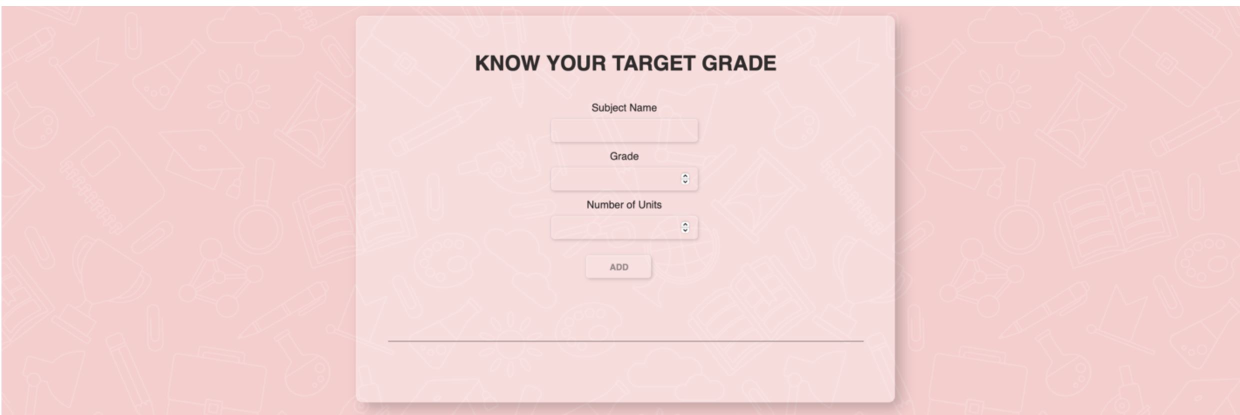
World Health Organization. (n.d.). *Cardiovascular diseases*. World Health Organization. Retrieved December 7, 2022, from https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1

V
APPENDICES

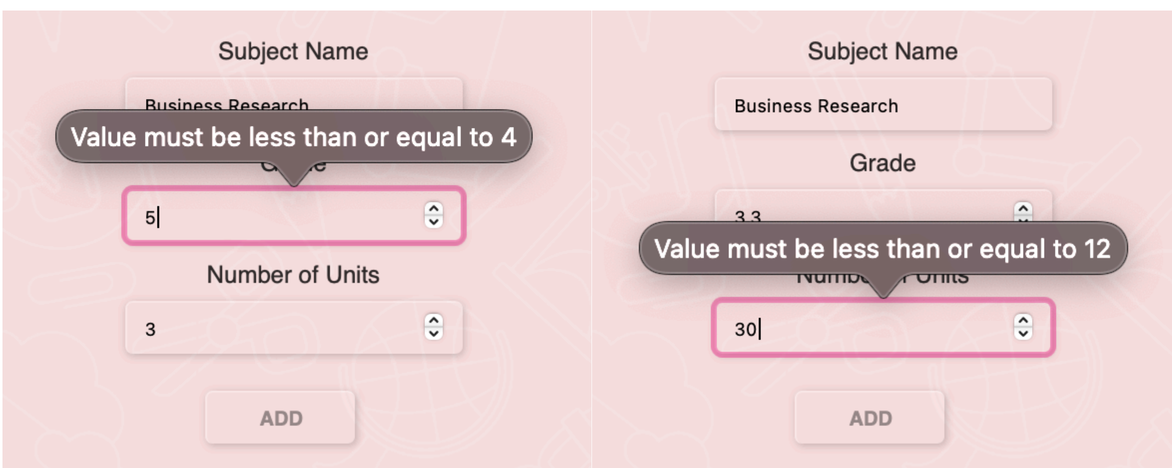
Appendix A. Grade Goal Seek Documentation

The following is a detailed explanation of how the goal seek program runs.

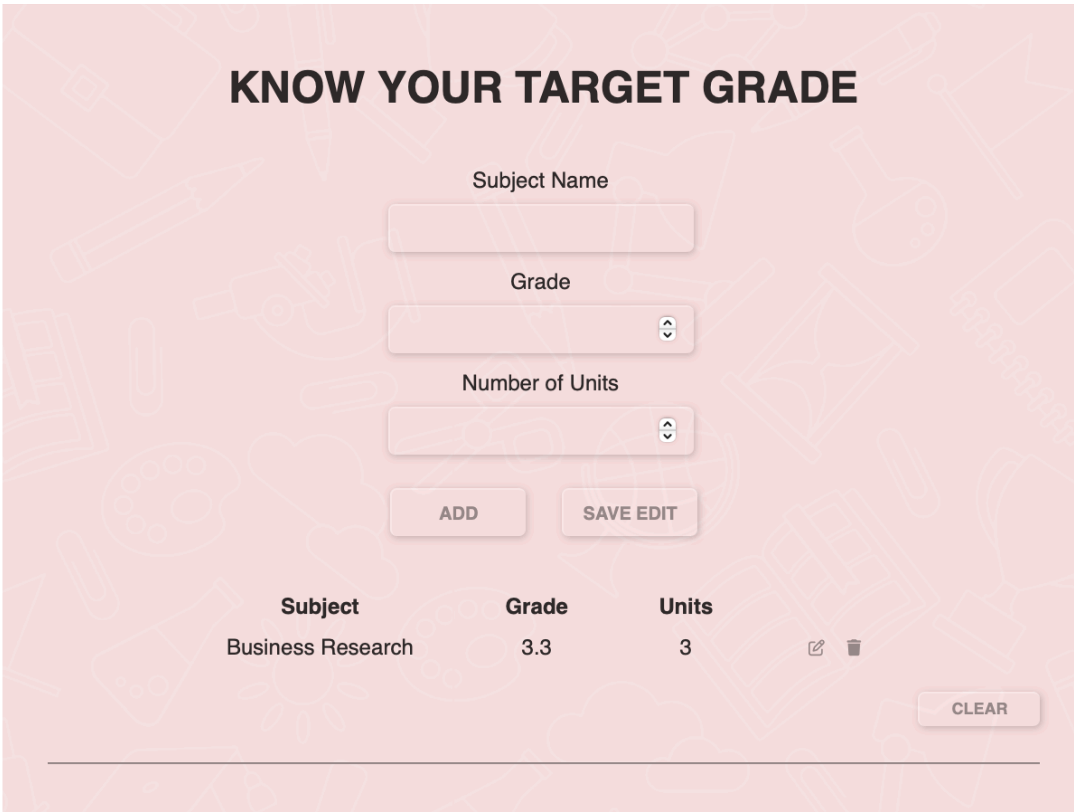
When the page is loaded, a form for inputting the subject name, grade, and number of units for that certain grade is shown. The subject name accepts text inputs while the grade and number of units accepts numerical inputs. The minimum numbers are set to zero while the maximum numbers are set to 4 and 12 for the grade and number of units, respectively. Only the grade input accepts decimal inputs.



When the add button is clicked and the values are invalid, an error message is prompted.

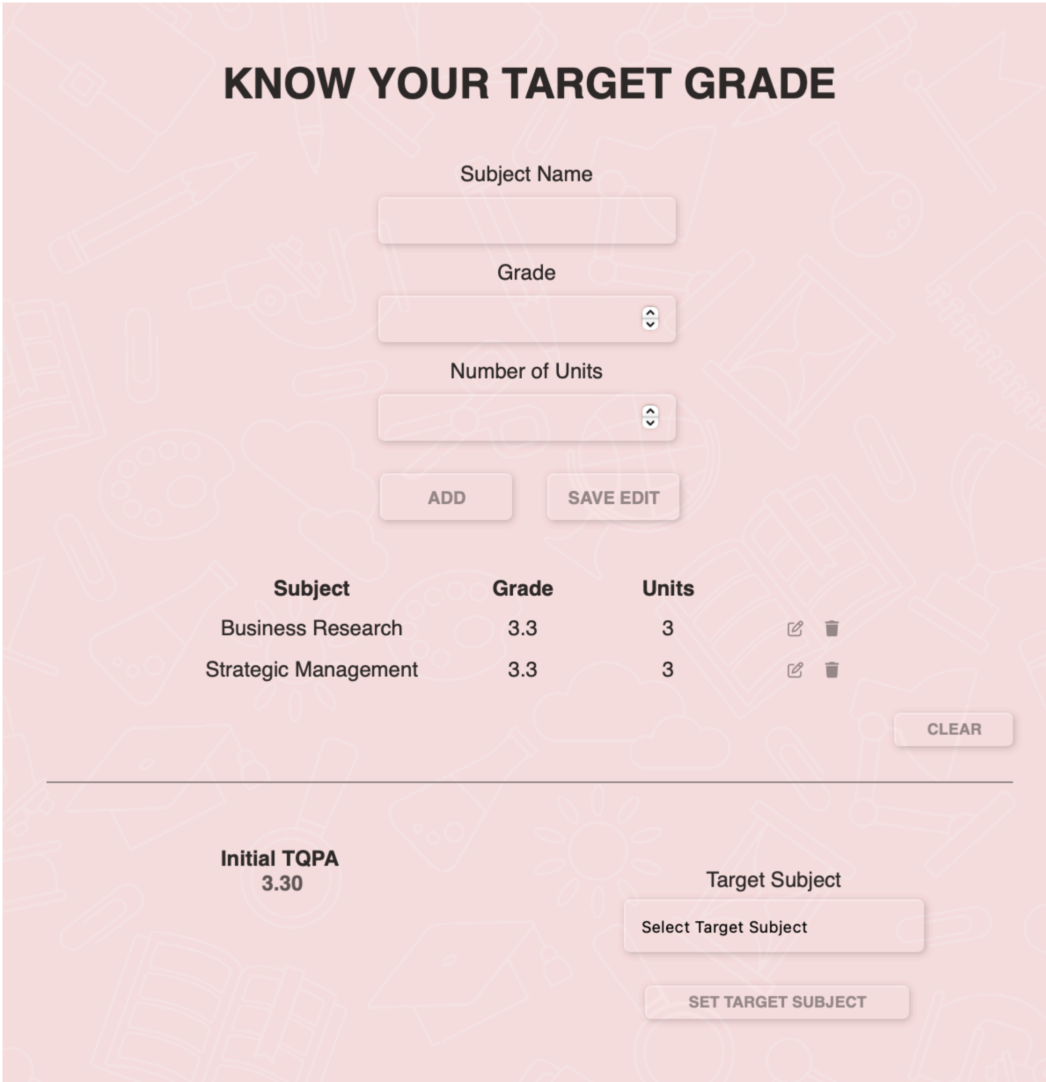


If the inputted data are valid, a row is inserted into the MySQL Database. No other operations can be performed unless more data is added.



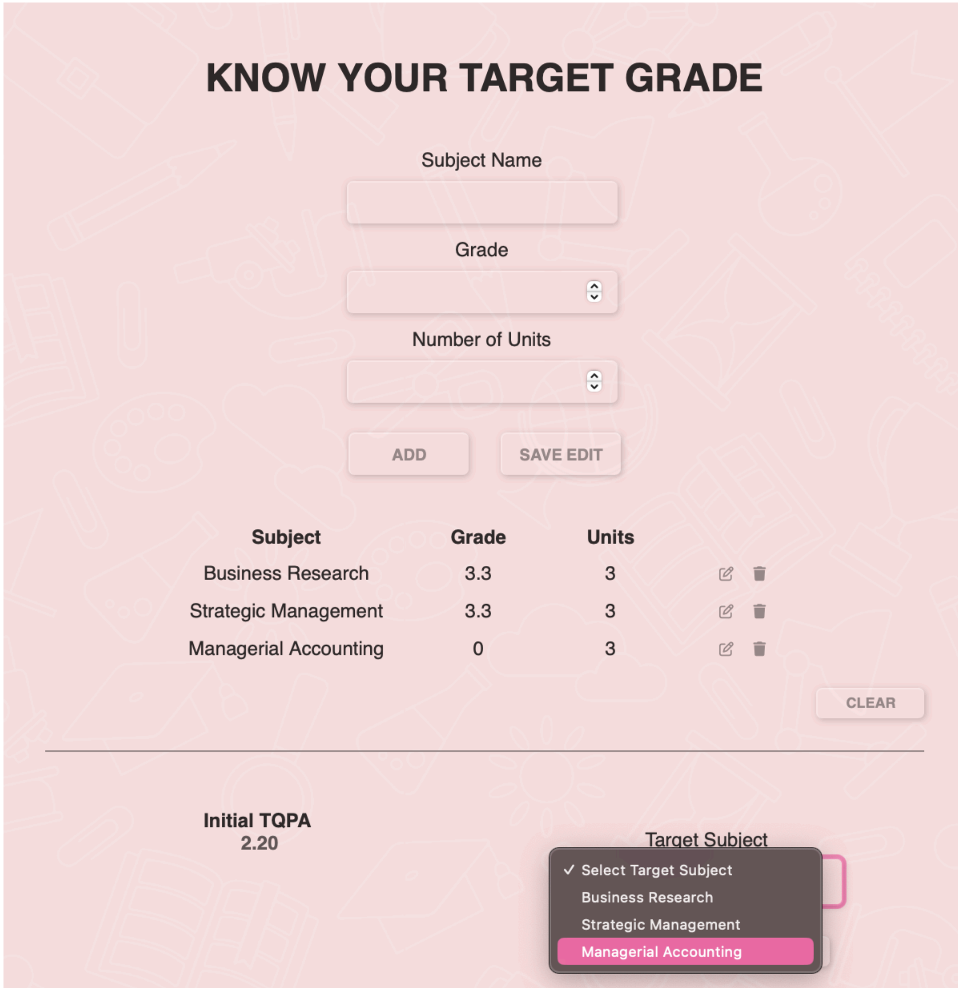
Once there are at least two rows of data, the initial TQPA is displayed and the user can finally have the option to select a target subject. This program does not accept letter grades the only option should be to set the grade to 0. Assuming there is no available target grade yet for the target subject, the grade can also be set to 0.

Users have the option to edit and delete rows of data using the edit and delete buttons displayed at the right hand side of the units.



Once all necessary data are inputted and the user is ready to set the target subject, they can click the select target subject option and select a subject where a list of all subjects are displayed.

When the target subject is set, more information is displayed. This includes the set target subject, the maximum achievable TQPA, and a form for setting a desired TQPA showing the initial TQPA as the minimum and a calculated maximum attainable TQPA as the maximum.

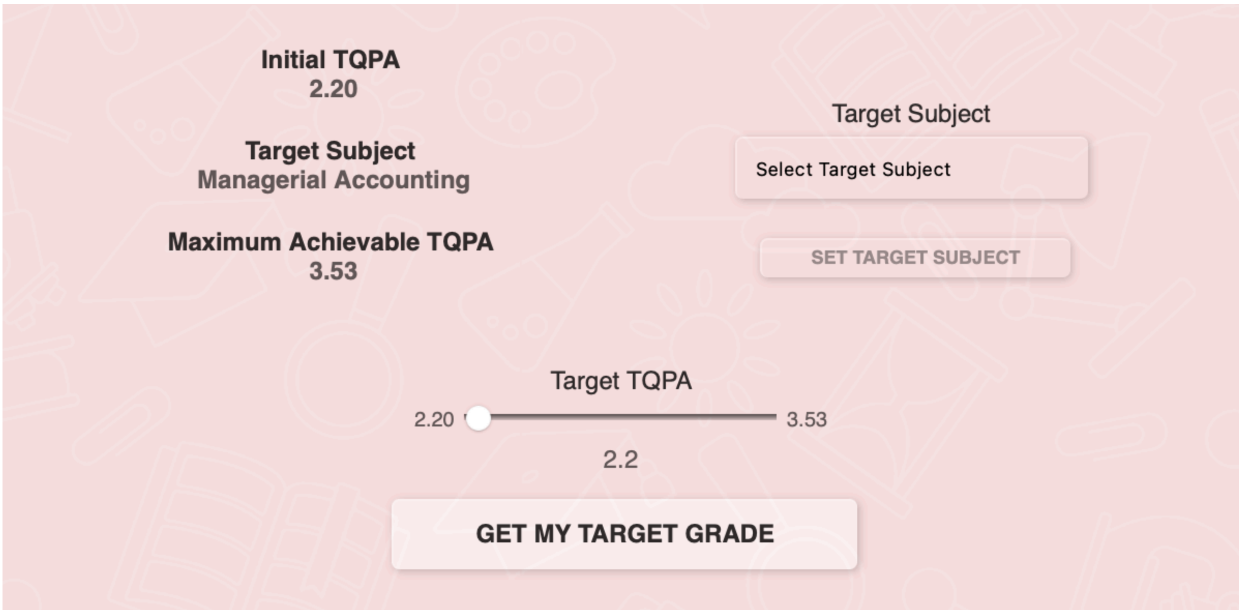


The following image shows the data in the MySQL Database before and after the target subject is set.

subject_id	subject_name	grade	units	istarget
1	Business Research	3.3	3	n
2	Strategic Management	3.3	3	n
3	Managerial Accounting	0	3	n

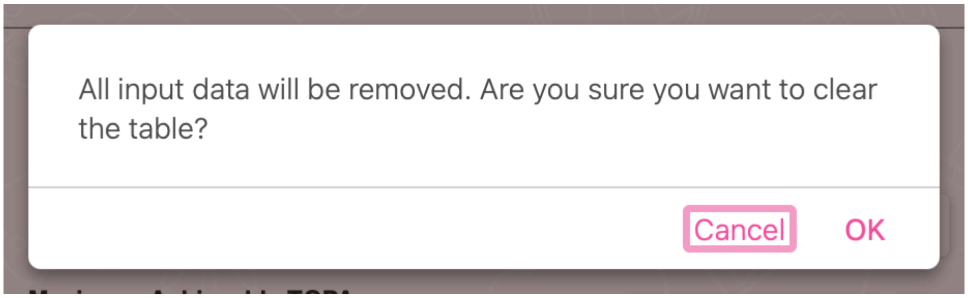
subject_id	subject_name	grade	units	istarget
1	Business Research	3.3	3	n
2	Strategic Management	3.3	3	n
3	Managerial Accounting	0	3	y

The maximum achievable TQPA ensures that the target subject grade does not exceed 4.0. However due to rounding off, numbers can be displayed as more than 4.0 with minimal decimal increments. This was calculated by iterating all the possible grade and TQPA combinations given the target subject and taking the TQPA where the target subject grade is 4.



Now, using the slider, the user can select a desired target TQPA. When the "GET MY TARGET GRADE" button is clicked, the selected target TQPA, the target subject name, and the target subject grade are selected.

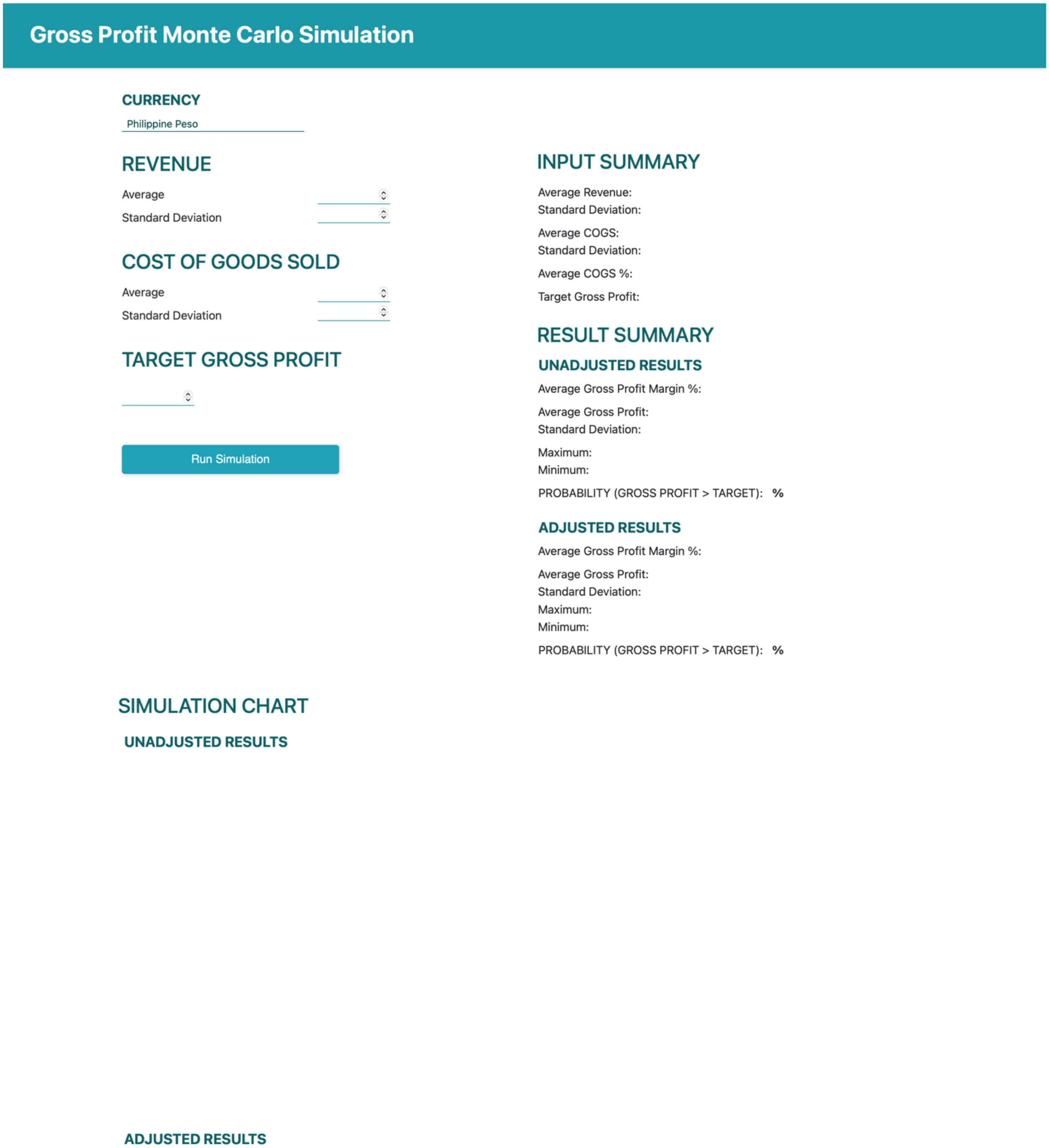
Once the user is done, the user has the option to clear all the inputted data using the clear button, modify the table, or keep it as it is for future references.



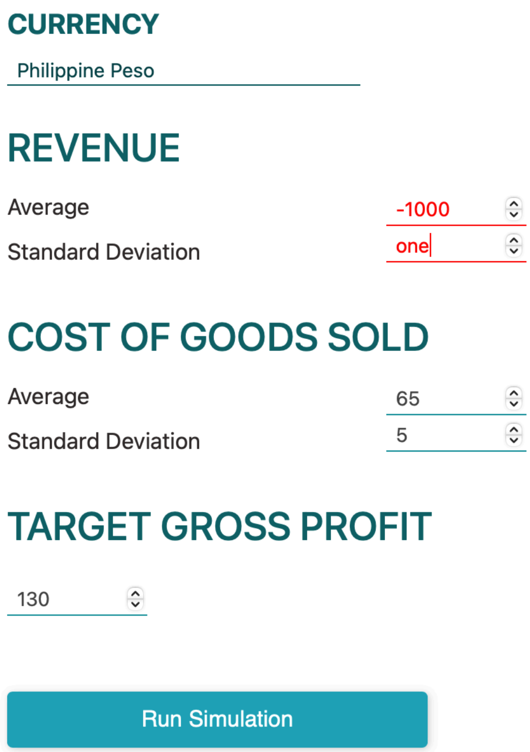
As a control measure, the user is asked to confirm when clearing the tables. In the MySQL database, this option truncates the data table.

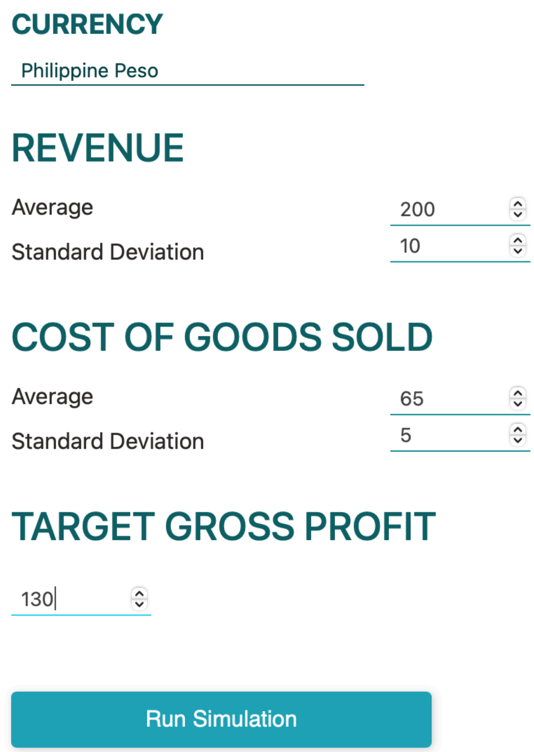
Appendix B. Gross Profit Monte Carlo Simulation Documentation

The Monte Carlo Simulation program is explained in detail as follows.

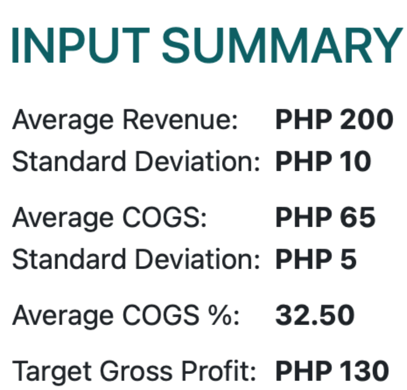


Once the web program is loaded, a blank form and results with empty values are displayed. The form includes the currency wherein the default currency is set to Philippine Peso(), average and standard deviation inputs for the revenue and cost of goods sold, and the target gross profit input. For the currency, users are given a list of worldwide currencies they can select from. Only non-negative numerical values are allowed as inputs under revenue, cost of goods sold, and target gross profit. As a control, the input box is designed to turn red to prompt the user if the value is invalid.





When the "Run Simulation" button is clicked, a JavaScript function that fills in the empty values in the Input Summary, Result Summary as well as the charts, is ran. The input summary displays the values added by the user. A calculation for the average cost of goods sold in percentage is added.



For the result summary, 1000 random values in a normal distribution for the revenue as well as the cost of goods sold are generated. An array of 1000 values for the gross profit was calculated by subtracting the individual values of cost of goods sold from the corresponding values under the revenues. The unadjusted results are calculations based on all the 1000 iterations. The adjusted results on the other hand, are calculations based only on the values that fall above the bottom 10% percentile and values that fall below the top 10% percentile.

RESULT SUMMARY

UNADJUSTED RESULTS

Average Gross Profit Margin %: **67.69**

Average Gross Profit: **PHP 135.39**
Standard Deviation: **PHP 10.92**

Maximum: **PHP 171**
Minimum: **PHP 94.15**

PROBABILITY (GROSS PROFIT > TARGET): **69.30%**

ADJUSTED RESULTS

Average Gross Profit Margin %: **67.24**

Average Gross Profit: **PHP 134.48**
Standard Deviation: **PHP 9.92**

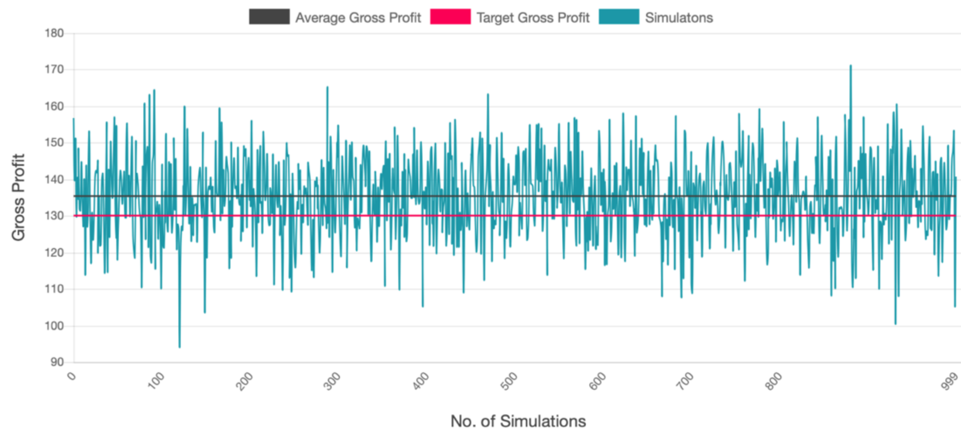
Maximum: **PHP 153.88**
Minimum: **PHP 103.73**

PROBABILITY (GROSS PROFIT > TARGET): **67.96%**

The adjusted and unadjusted results are also presented under Simulation Charts in separate line graphs for comparison. The unadjusted chart features all the 1000 gross profit values (0-999). Meanwhile, the adjusted chart only features the values that fall within the calculated range.

SIMULATION CHARTS

UNADJUSTED RESULTS

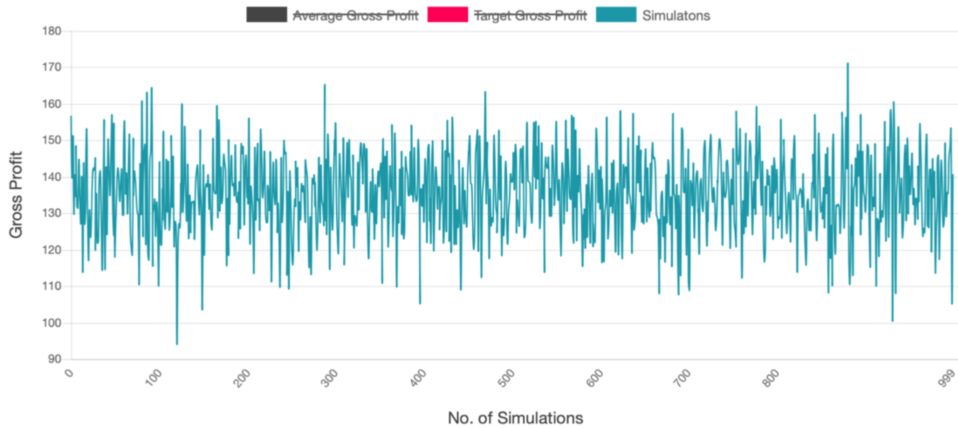


ADJUSTED RESULTS

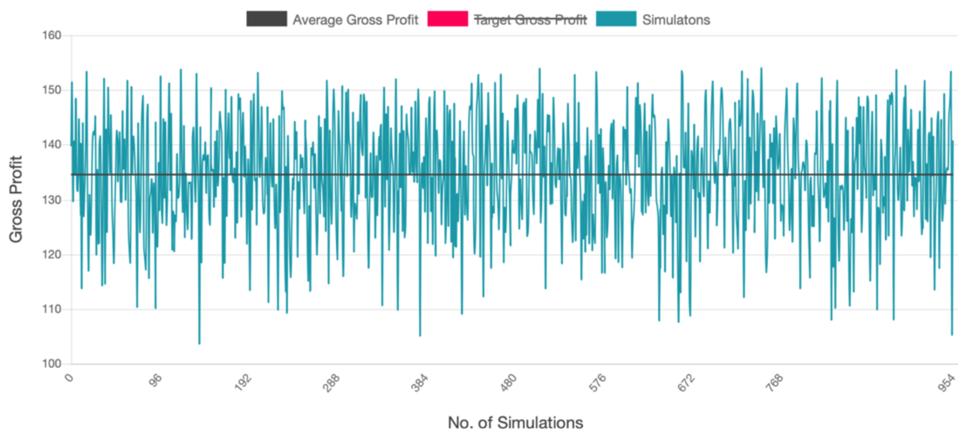
Both charts are dynamic in a way that the user can remove the lines they wish to exclude in their analysis. The user can also regenerate new values and have them calculated every time the "Run Simulation" button is clicked.

SIMULATION CHARTS

UNADJUSTED RESULTS



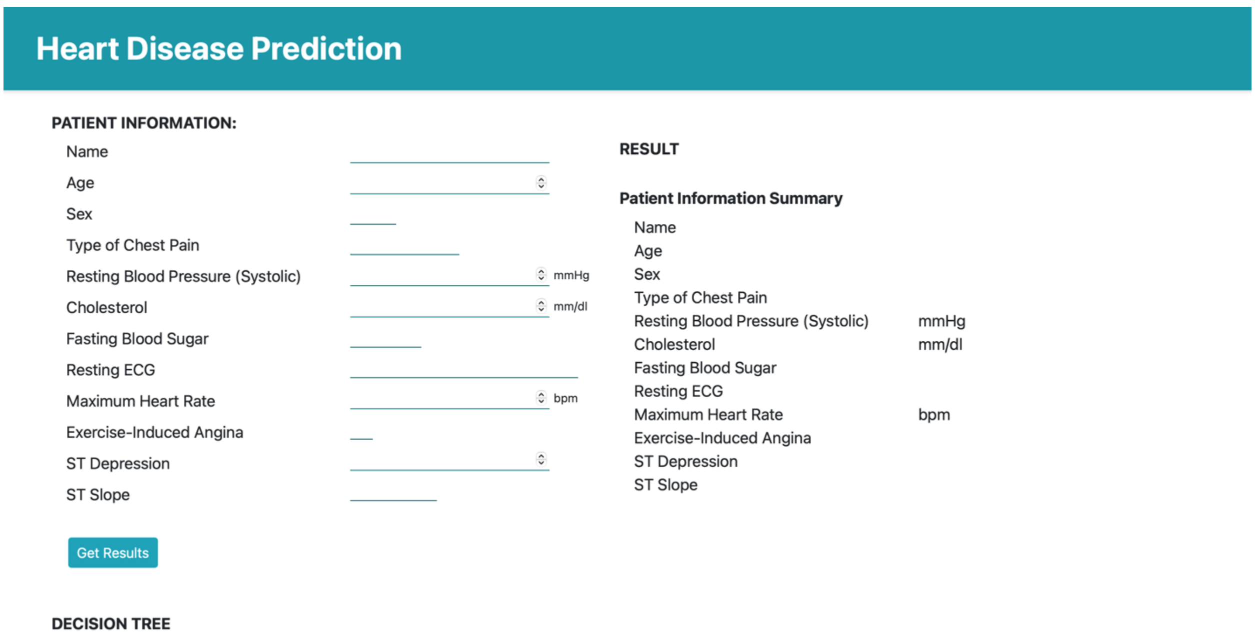
ADJUSTED RESULTS



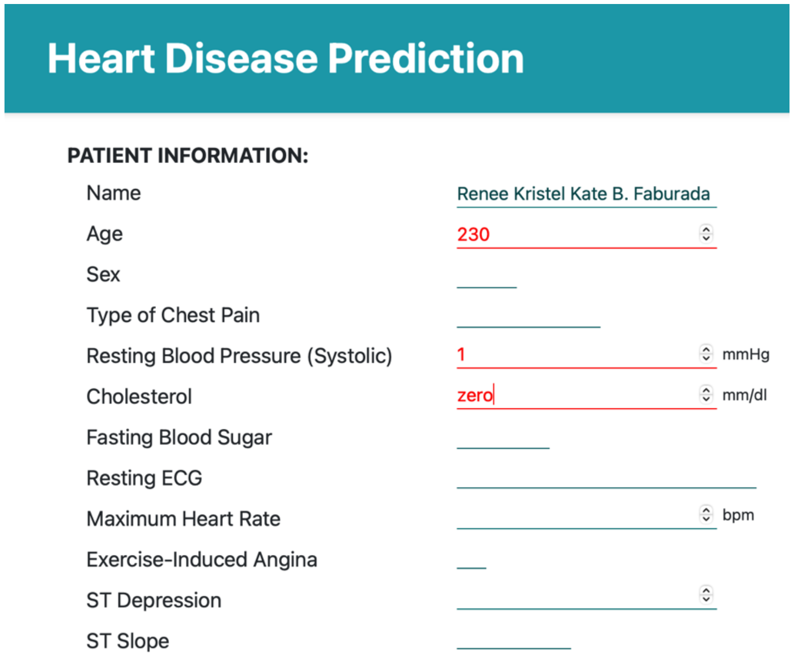
Appendix C. Heart Disease Prediction Decision Tree Documentation

The following explains in detail how the decision tree prediction web program runs.

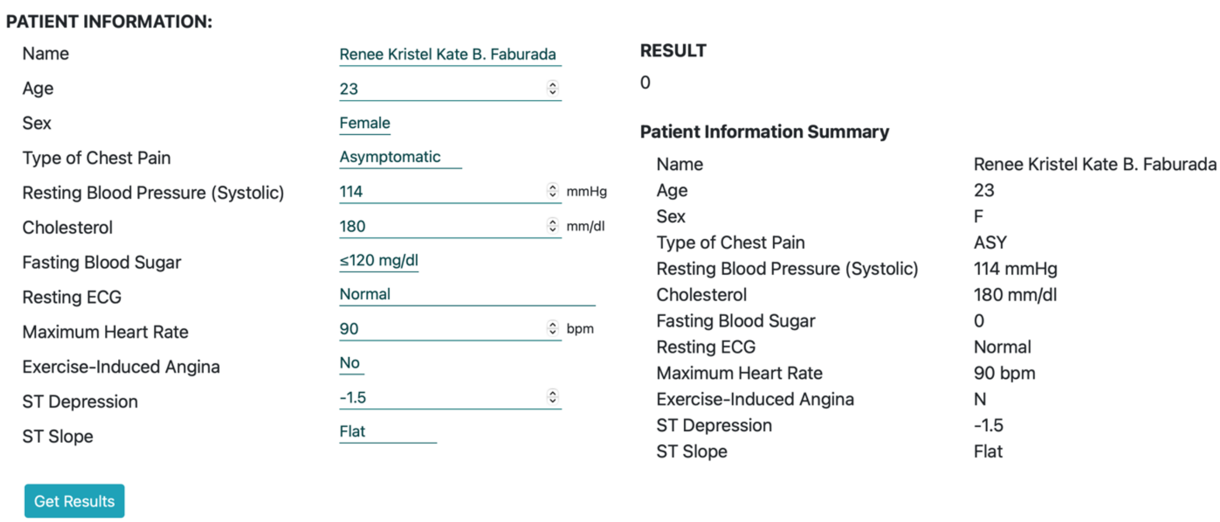
As the web program is loaded, the user is given a form requiring the relevant information. This includes the patient name and the 11 dependent variables used in the dataset namely the age, sex, type of chest pain, resting systolic blood pressure in mmHG, cholesterol in mm/dl, fasting blood sugar, resting ECG result, maximum heart-rate achieved, whether or not the patient experiences exercise induced angina, the old peak or the ST depression, and the ST slope. The information about the data has been presented in *Figure 4.*



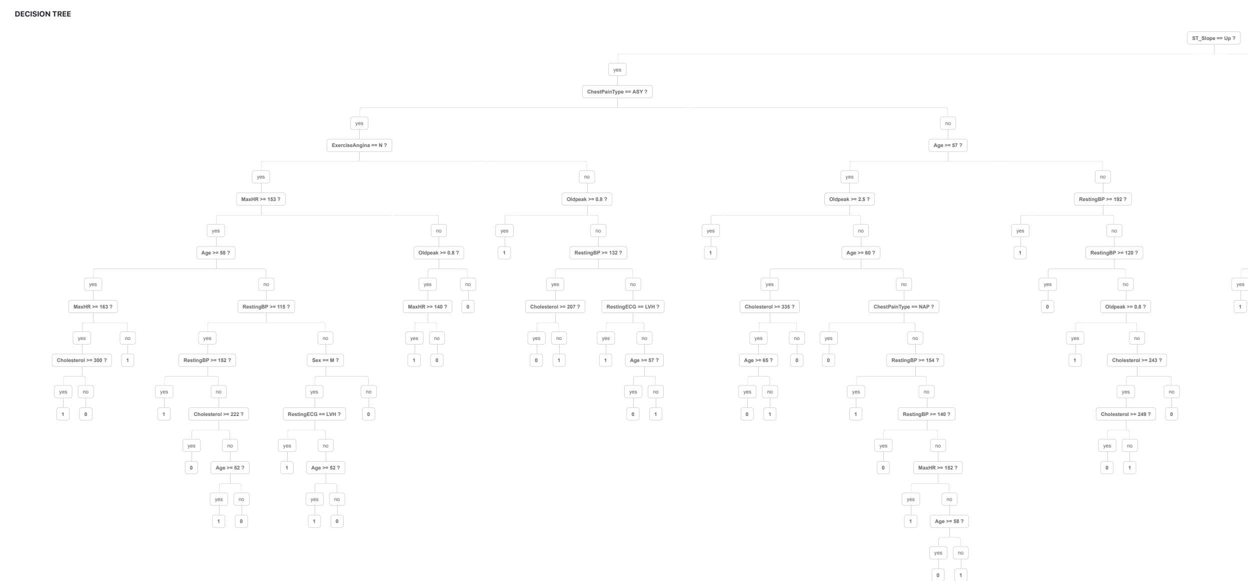
If data does not align with that of the required type as described in *Figure 4,* it is deemed invalid and the user is prompted as the input block color changes to red.

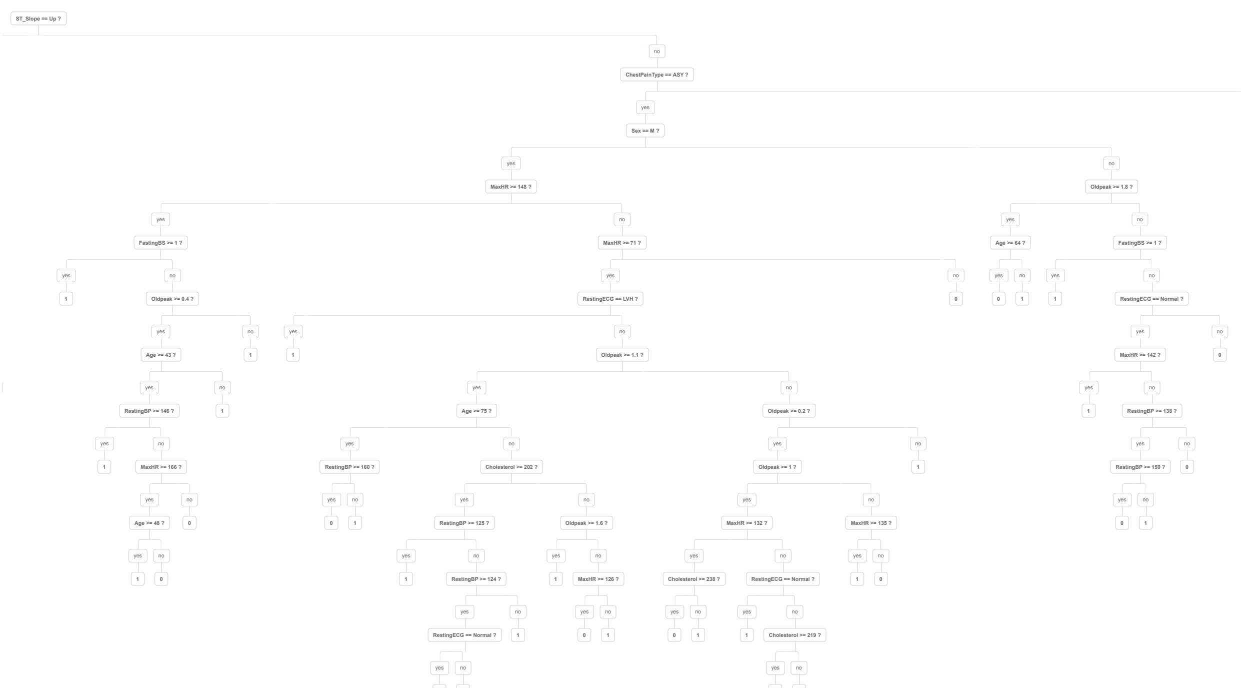


Once the user has clicked the "Get Results" button, a JavaScript function for viewing the results, along with the patient information summary is displayed. A result of 0 means that it is likely that the patient does not have a heart disease while a result of 1 means that it is likely that the patient has a heart disease.

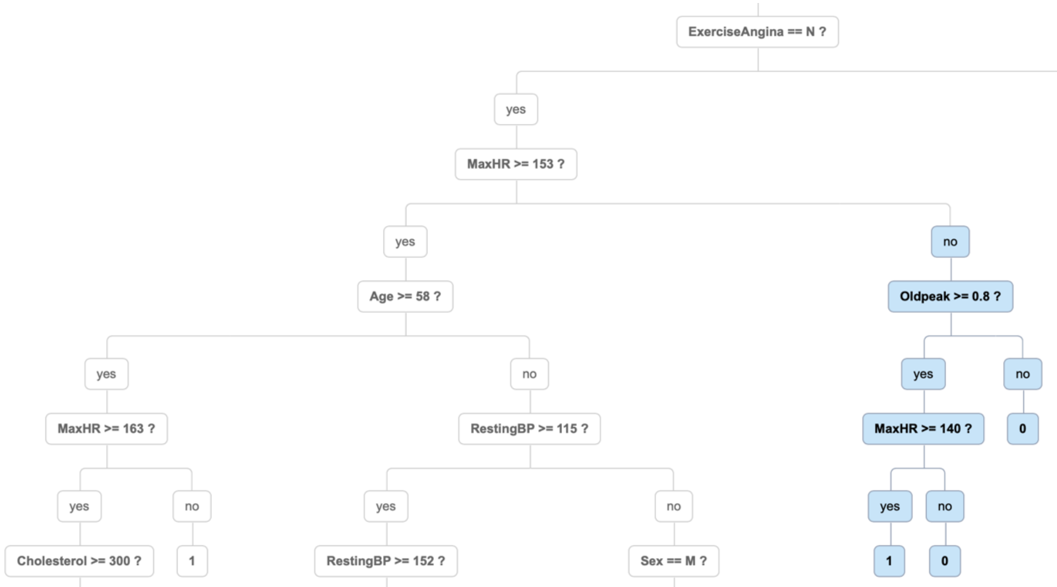


The web program also displays the decision tree built using a JavaScript function where the user can manually track and compare their data.





The user can easily track the nodes as the color changes to blue when hovered over.



Appendix D. Source Code

The source code of the three projects can be found at the following GitHub link:

https://github.com/reneekristelkate/SMProject.

An index page is included for ease of navigation.

Appendix E. Curriculum Vitae

Renee Kristel Kate B. Faburada

E. Buling Street, Balugo, Dumaguete City, Negros Oriental 6200
reneebfaburada@su.edu.ph
+63 921 677 5633

EDUCATION

Silliman University — Dumaguete City
Bachelor of Science in Business Administration — Expected June 2022
Major in Business Analytics

Related Coursework: Database Management Systems; Data Warehousing; Analytics; Enterprise Resource Planning; Financial, Project, and Strategic Management; Programming

Saint Louis School of Don Bosco, Inc. — Dumaguete City
Senior High School, With High Honors — March 2018

Saint Louis School of Don Bosco, Inc. — Dumaguete City
Junior High School, 2nd in Excellence — March 2016

West City Science Elementary School — Dumaguete City
Elementary — March 2012

EXPERIENCE

Silliman University Library — Dumaguete City
Student Library Assistant — 2019

- Worked at the Circulation and Filipiniana sections, lending and receiving library-owned materials
- Assisted students, staff, and visitors, referring them to appropriate resources
- Identified valuable articles and prepared clippings from newspapers for circulation
- Distributed resources to the different departments in the university
- Ensured the cleanliness and organization in the library
- Maintained the physical quality of books and other library resources
- Answered incoming phone calls

SKILLS

Languages: English, Cebuano, Filipino
Software: PyCharm, Visual Studio Code, Mac OS, Windows OS, Microsoft Word, Microsoft Excel, Microsoft PowerPoint, Pages, Numbers, Keynote, Krita
Programming: SQL, HTML, CSS, Python, PHP, JavaScript
Interpersonal Skills: Attention to Detail, Verbal and Written Communication, Time Management, Adaptability, Having Initiative, Organizing, Quick Learning Skills, Good Work Ethic
Interests: Digital Art, Graphic Design, Reading

AWARDS & ACTIVITIES

Silliman University College of Business Administration COMELEC — Dumaguete City
Member — 2021-2022

Silliman University Business Analytics Students' Society — Dumaguete City

Auditor 2021-2022

Saint Louis School of Don Bosco, Inc. Bosconian Leap (Student Publication) Dumaguete City
Editor-in-Chief, English Section 2016-2018

Saint Louis School of Don Bosco, Inc. Bosconian Year Book Dumaguete City
Editor-in-Chief, English Section 2016-2018

National Schools Press Conference Pagadian City
Participant, English News Writing 2017

Regional Schools Press Conference Bacolod City
3rd Place, English News Writing 2017

Division Schools Press Conference Dumaguete City
2rd Place, English News Writing 2017

Journalism Award in English News Writing Dumaguete City
Saint Louis School of Don Bosco, Inc. 2016

GRANTS & SCHOLARSHIPS

Honor Scholarship Grant Dumaguete City
Silliman University, College of Business Administration 2018–2022

CERTIFICATIONS

Bloomberg Certified Financial Markets Professional Dumaguete City
Silliman University June 22, 2022

Google Analytics for Beginners
Certificate of Completion December 2022

Advanced Google Analytics
Certificate of Completion December 2022