ANALYSIS AND DETERMINATION OF FACTORS ASSOCIATED WITH STUDENT PERFORMANCE BASED ON COMPOSITE ASSESSMENT SCORES

Suhaila Bahrom, Balqis Hisham

Centre for Foundation Studies, International Islamic University Malaysia,

*Author's email address: *suhaila_b@iium.edu.my*

Abstract: Examining final scores among pre-university engineering students is crucial for understanding their academic performance and identifying factors contributing to success or challenge in their educational journey. These scores are pivotal indicators of students' grasp of fundamental engineering principles and readiness for higher education. This study examines the correlation between different course assessments and final exam scores in a Mathematics course for pre-university engineering students. A dataset comprising assessments such as quizzes, open-book tests, and tutorials was collected from 552 pre-university engineering students at the Centre for Foundation Studies, International Islamic University Malaysia, for the 2023/2024 cohort. Regression analysis was employed to identify the significance course assessments, which were carried out using Python. The study revealed that all the quizzes, including open book test 2 and open book test 3, are significantly correlated with final examination scores with an adjusted R-squared of 0.467. This value indicates that 46.7% of the variation in final examination scores can be predicted by combining all quizzes and two open-book tests. This study examines the effectiveness of course assessments in predicting the final examination performance of students in pre-university engineering programs. Furthermore, it presents valuable recommendations for enhancing assessment strategies to support and foster student achievement more effectively.

Keywords: Regression analysis; Pre-University; Python; Mathematics

1. INTRODUCTION

1.1 Background

Course assessment is crucial to measure students' understanding of the course material. Formative assessment monitors students' learning progress during a program. Its objective is to provide continuous feedback to students and instructors to identify strengths, weaknesses, and areas for improvement (Lee et al., 2020). Examples of formative assessments are quizzes, tests, take-home exercises, group projects, and case studies (Granberg et al., 2021). Different course has different methods to assess the students depending on the course learning outcome. The Assessment for Learning model emphasizes that assessments should be used to enhance learning and provide feedback that supports student development (Black & Wiliam, 1998). Educators should implement suitable assessments to ensure students understand the course content (Weldmeskel & Michael, 2016). In addition, the Theory of Formative Assessment agrees with the idea that using different types of assessments throughout the learning process allows teachers to monitor how well students are learning (Wiliam, 2011). Based on this feedback, teachers can adjust their teaching methods to support student learning and improve educational outcomes.

1.2 Literature Review

This literature review explores studies that employ multiple linear regression (MLR) to predict final exam scores based on various assessment methods.

Tutorials

Tutorials (Tutor), often structured as supplementary exercises to reinforce classroom learning, have been studied extensively for their predictive value in academic outcomes. In this study context, usually, students will be given a set of 20 to 25 questions from each chapter. Research by Smith (2017) found that tutorial attendance and participation positively correlate with higher exam scores, highlighting the role of active learning strategies in enhancing student performance. Similarly, Lee and Jones (2019) demonstrated through MLR analysis that tutorial engagement significantly predicts final exam outcomes across multiple disciplines, underscoring the importance of personalized academic support.

Quizzes

Quizzes (Q) serve as formative assessments that measure students' understanding of course material. In this study, students need to take 3 quizzes in 1 semester. The structured questions involve 3 levels: easy, moderate, and difficult. Studies by Brown et al. (2022) have shown that quiz performance when integrated into predictive models using MLR, provides early indicators of exam success.

Open Book Tests

Open book tests (OBT) represent a formative assessment, where students can refer to their notes and books to answer a set of questions. In this study context, students need to take 3 open-book tests per semester. Research by Garcia and Smith (2018) explored how open book test scores, incorporated into MLR models, contribute to predicting final exam performance. Their findings suggest that students who excel in open book tests often demonstrate higher-order cognitive skills that translate into improved exam results, highlighting the pedagogical benefits of this assessment format.

1.3 Research Questions

In this research paper, there are 2 research questions:
1. Is there any correlation between different course assessments and final exam scores in Mathematics course?
2. What are the significance assessments in the Mathematics course?

1.4 Objectives

In this research paper, there are 2 objectives:
1. To examine the correlation between different course assessments and final exam scores in a Mathematics course for pre-university engineering students.
2. To identify the significance assessments in the Mathematics course.

1.5 Significance of the study

This study's significance was to determine the potential approach to transform assessment practices at the Centre for Foundation Studies, IIUM, by providing a comprehensive analysis of how different types of assessments, such as tutorials, quizzes, and open-book tests, affect student performance on the final exam scores. This study provides valuable insights for educators to enhance assessment methods, curriculum design, and student performance.

2. METHODOLOGY

2.1 Data collection

The data has been collected from the pre-university engineering students' results for Mathematics I in the Centre for Foundation Studies, IIUM, for cohort 2023/2024. There are 552 students from the foundation of engineering programs are included in this study.

2.2 Data Preprocessing

Data preprocessing is a crucial step in data analysis. In this analysis, data preprocessing has been used to improve the data's quality and reliability by addressing missing values, outliers, and inconsistent formats. Figure 2 shows a heatmap of missing data for the dataset used in this study. The heatmap in Figure 1 shows a uniform colour, indicating that there is no missing data across all variables and samples in the dataset.

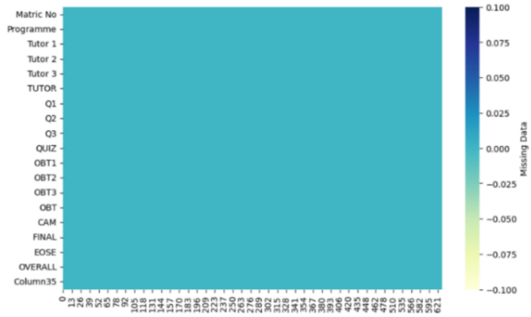


Fig. 1. Diagnostic Plots for Regression Model

3. RESULTS AND DISCUSSIONS

3.1 Data Descriptive

This analysis includes nine course assessments as independent variables: Tutor 1, Tutor 2, Tutor 3, Q1, Q2, Q3, OBT1, OBT2, and OBT3. The dependent variable is the final examination score. Figure 2 illustrates the summary of the results. The highest examination score is 67 out of 70, while the lowest is 0. The average score on the final exam was 40.3. The final exam score distribution is illustrated in Figure 3, showing a normal distribution.

	Tutor 1	Tutor 2	Tutor 3	Q1	Q2	Q3	OBT1	OBT2	OBT3	FINAL
count	552.0000	552.0000	552.0000	552.0000	552.0000	552.0000	552.0000	552.0000	552.0000	552.0000
mean	4.9746	4.9493	4.9230	10.8859	14.4783	16.0054	7.2989	9.6214	9.7428	40.3043
std	0.2682	0.4110	0.5595	4.3578	3.9630	2.8178	3.1438	3.7373	3.0517	10.7650
min	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000
25%	5.0000	5.0000	5.0000	8.0000	12.0000	15.0000	5.0000	7.0000	8.0000	33.0000
50%	5.0000	5.0000	5.0000	11.0000	15.0000	16.0000	7.0000	10.0000	10.0000	41.0000
75%	5.0000	5.0000	5.0000	14.0000	17.0000	18.0000	10.0000	13.0000	12.0000	47.0000
max	5.0000	5.0000	5.0000	20.0000	20.0000	20.0000	15.0000	15.0000	15.0000	67.0000

Fig. 2. Descriptive Summary of Variables

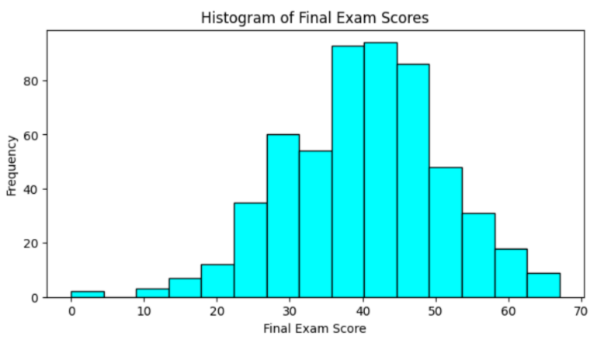


Fig. 3. Distribution of Final Exam Scores

3.2 Multiple Linear Regression Analysis

The study aimed to identify the significant formative assessments that can be used to predict students' final exam scores through linearity tests. Figure 4 illustrates the results of multiple linear regression analysis. The F-statistic is significant ($p < 0.001$), suggesting the model fits well. Among the predictors, Q1, Q2, Q3, OBT2, and OBT3 have statistically significant coefficients $\left(p - value < 0.001\right)$, indicating they are significant predictors of the dependent variable. In contrast, the coefficients for Tutor 1, Tutor 2, Tutor 3, and OBT1 are not statistically significant, suggesting their impact on the final exam score is negligible within this model $\left(p - value > 0.05\right)$. Therefore, the next step, which is model selection, is important to ensure that only significant predictors are considered in the final model. The hypothesis for the linear relationship test is as follows:

H_0 : Neither of the independent variables is related to the independent variables.

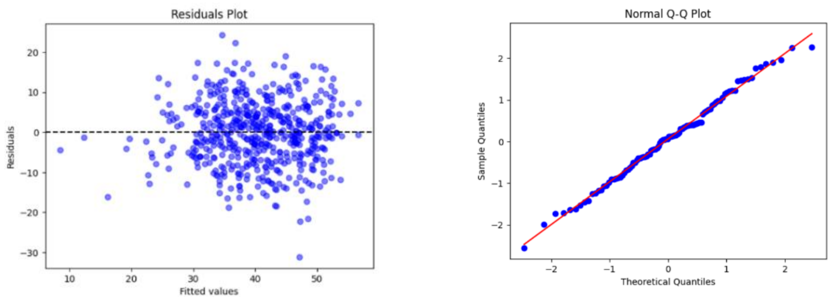
H_0 : At least one of the independent variables is related to the dependent variables.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                  FINAL   R-squared:                       0.474
Model:                            OLS   Adj. R-squared:                  0.466
Method:                 Least Squares   F-statistic:                     54.37
Date:                Tue, 16 Jul 2024   Prob (F-statistic):           4.11e-70
Time:                        06:46:33   Log-Likelihood:                -1916.9
No. Observations:                 552   AIC:                             3854.
Df Residuals:                     542   BIC:                             3897.
Df Model:                           9
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -3.9130      6.491     -0.603      0.547     -16.664       8.839
Tutor 1        1.0956      1.437      0.762      0.446      -1.728       3.919
Tutor 2        0.6277      1.011      0.621      0.535      -1.358       2.614
Tutor 3        0.1932      0.740      0.261      0.794      -1.260       1.647
Q1             0.2804      0.090      3.100      0.002       0.103       0.458
Q2             0.7277      0.103      7.091      0.000       0.526       0.929
Q3             0.4184      0.139      3.013      0.003       0.146       0.691
OBT1           0.0336      0.126      0.267      0.790      -0.213       0.280
OBT2           0.5654      0.105      5.365      0.000       0.358       0.772
OBT3           0.8970      0.129      6.977      0.000       0.644       1.150
==============================================================================
Omnibus:                        1.567   Durbin-Watson:                   1.841
Prob(Omnibus):                  0.457   Jarque-Bera (JB):                1.368
Skew:                          -0.094   Prob(JB):                        0.505
Kurtosis:                       3.155   Cond. No.                         600.
==============================================================================
```

Fig. 4. Results of OLS Regression Analysis

3.3 Residuals Analysis

Residual analysis is a method used in regression analysis to assess a model's goodness of fit and verify the underlying assumptions. It involves checking for linearity, homoscedasticity, normality, and independence (Martin et al., 2017). Figure 5 depicts the diagnostic plots for the regression model. Figure 5(a) shows that the residuals appear randomly scattered around the horizontal axis, suggesting that the model's assumptions of linearity and homoscedasticity are reasonably met. The Normal Q-Q Plot in Figure 5(b) shows the points close to the diagonal line, indicating that the residuals are approximately normally distributed. These plots suggest that the regression model is well-fitted and its assumptions are valid. The residuals' distribution can be clearly observed in Figure 6, illustrating their normal distribution.



(a) Residual plot (b) Normal Q-Q plot

Fig. 5. Diagnostic Plots for Regression Model

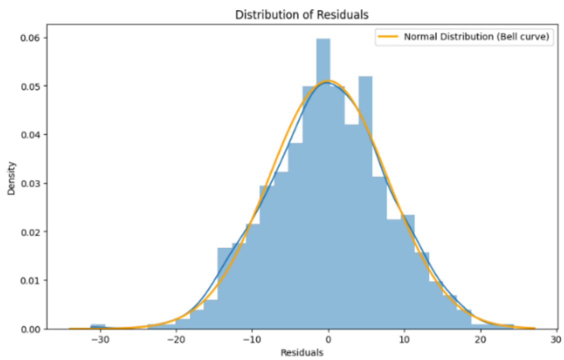


Fig. 6. Distribution of Residuals

3.4 Multicollinearity

Multicollinearity is a statistical concept that occurs when two or more independent variables in a model are highly correlated (Kim, 2019). In a regression model, an independent variable can be predicted from another independent variable. Figure 3 shows a correlation heatmap, visualizing the relationships between dependent and independent variables using colours representing strength and direction. Lighter colours indicate stronger positive correlations, and darker colours indicate stronger negative correlations (W. Ding et al., 2023). The correlation heatmap indicates that the independent variables generally have low correlation coefficients, with most values below 0.4. This suggests a low degree of linear association between the independent variables. There are no issues with multicollinearity in this dataset.

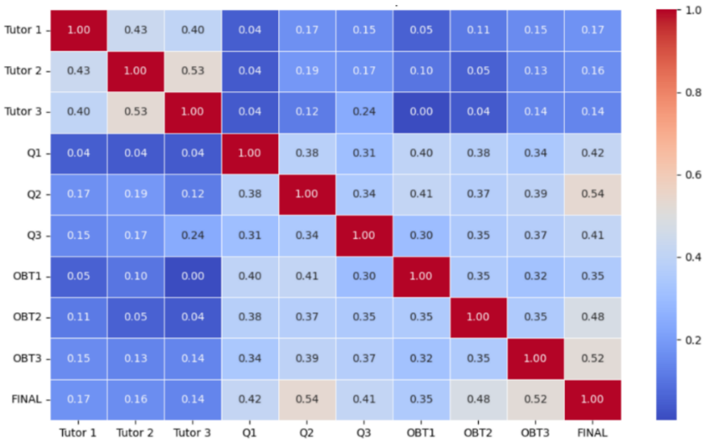


Fig. 7. Correlation heatmap

	Variable	VIF
0	const	375.711756
1	Tutor 1	1.322931
2	Tutor 2	1.536582
3	Tutor 3	1.525699
4	Q1	1.382894
5	Q2	1.472219
6	Q3	1.362545
7	OBT1	1.389073
8	OBT2	1.380563
9	OBT3	1.369864

Fig. 8. Variance Inflation Factor (VIF) Values

The VIF values for different variables in a regression model to assess multicollinearity are displayed in Figure 8. VIF is a measure used to determine the degree of multicollinearity between independent variables in a regression model (O'brien, 2007). A high VIF value indicates high multicollinearity with other variables. All predictors, have VIF values ranging from 1.322931 to 1.536582. These values are well below the common threshold of 10, indicating that multicollinearity is not a significant issue for these predictors in the model (Salmerón Gómez et al., 2020).

3.5 Model Selection

Model selection in regression analysis involves selecting the most significant subset of predictors contributing to the dependent variable (J. Ding et al., 2018). This process is essential to ensure the model is simple and easy to understand, avoiding overfitting and underfitting. The output of the selection model is depicted in Figure 9.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                  FINAL   R-squared:                       0.472
Model:                            OLS   Adj. R-squared:                  0.467
Method:                 Least Squares   F-statistic:                     97.68
Date:                Fri, 12 Jul 2024   Prob (F-statistic):           1.96e-73
Time:                        21:14:46   Log-Likelihood:                -1918.1
No. Observations:                 552   AIC:                             3848.
Df Residuals:                     546   BIC:                             3874.
Df Model:                           5
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          4.7937      2.036      2.354      0.019       0.794       8.794
Q1             0.2756      0.088      3.129      0.002       0.103       0.449
Q2             0.7541      0.099      7.612      0.000       0.559       0.949
Q3             0.4508      0.135      3.341      0.001       0.186       0.716
OBT2           0.5663      0.104      5.453      0.000       0.362       0.770
OBT3           0.9164      0.127      7.193      0.000       0.666       1.167
==============================================================================
Omnibus:                        1.590   Durbin-Watson:                   1.835
Prob(Omnibus):                  0.452   Jarque-Bera (JB):                1.390
Skew:                          -0.095   Prob(JB):                        0.499
Kurtosis:                       3.156   Cond. No.                         173.
==============================================================================
```

Fig. 9. Results of Regression Analysis with Significant Predictors

Based on Figure 9, the OLS regression results for predicting the final exam score include the predictors Q1, Q2, Q3, OBT2, and OBT3. The R-squared adjusted is 0.467, indicating 46.7% variation in the final exam score can be explained by the predictors. The F-statistic of 97.68 is highly significant ($p < 0.0001$), indicating that the model fits well. All predictor variables have positive and significant coefficients, suggesting that increases in Q1, Q2, Q3, OBT2, and OBT3 are associated with increases in the final exam score. The predictive model for the final exam score can be written as Equation (1):

$$\hat{y}_{FINAL} = 4.7937 + 0.2756_{Q1} + 0.7541_{Q2} + 0.4508_{Q3} + 0.5663_{OBT2} + 0.9164_{OBT3} \quad (1)$$

4. CONCLUSIONS

Based on the findings from the multiple linear regression (MLR) and the correlation analysis, several conclusions can be drawn regarding factors influencing student performance in Mathematics course based on composite assessment scores. The MLR analysis in this study concluded that some of the composite assessments, which are Q1, Q2, Q3, OBT2, and OBT3, can be used to predict the final exam scores where the p-value of these assessment methods has a statistically significant coefficient. This highlights the importance of these assessments in assessing and potentially improving students' performance. Whereas for Tutor 1, Tutor 2, Tutor 3, and OBT1, they are not statistically significant. This might be due to the questions from tutorials being discussed in class with the lecturers. Therefore, students have ample time to ask questions and make corrections. This is why a majority of students are able to secure perfect scores in Tutorials 1, 2, and 3. Hence, these dependent variables are not a good indicator for predicting final exam scores. Based on the correlation analysis, it gives additional insights to find the relationship between all the independent variables with the dependent variable, which is the final exam scores. The result shows a strong and positive correlation between final exam scores and Q2, OBT2, and OBT3. It suggested that these assessment methods are associated with students' overall performance. Conversely, weaker correlations were observed for other variables, such as Tutor 1, Tutor 2, Tutor 3, and OBT1, indicating their minimal influence on the final exam scores within the context of this study.

Overall, a model developed in this study effectively explains 46.7% of the variance in the final exam scores, indicating a moderate level of predictive accuracy. These findings emphasize the significance of specific formative assessments in predicting and potentially enhancing student outcomes in Mathematics course. Educators can modify instructional strategies and implementation aimed at improving students' academic excellence. By focusing on assessments that show stronger correlations with final scores, educational practices can be refined to support better student learning and achievement in mathematics and potentially other subjects.

5. ACKNOWLEDGEMENT

The researchers would like to thank the Centre for Foundation Studies, IIUM, for their support and resources throughout this study.

REFERENCES

Black, P., & Wiliam, D. (1998). “Assessment and Classroom Learning.” *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.

Brown, A., & Smith, B. (2022). Predicting student exam performance using multiple linear regression: A meta-analysis. *Journal of Educational Research*, *45*(2), 201-215.

Ding, J., Tarokh, V., & Yang, Y. (2018). Model Selection Techniques: An Overview. *IEEE Signal Processing Magazine*, *35*(6), 16–34. https://doi.org/10.1109/MSP.2018.2867638

Ding, W., Goldberg, D., & Zhou, W. (2023). PyComplexHeatmap: A Python package to visualize multimodal genomics data. *IMeta*, *2*(3). https://doi.org/10.1002/imt2.115

Garcia, M., & Smith, K. (2018). Open book tests and their impact on exam performance: A case study. *Journal of Applied Educational Studies, 28*(3), 321-335.

Granberg, C., Palm, T., & Palmberg, B. (2021). A case study of a formative assessment practice and the effects on students' self-regulated learning. *Studies in Educational Evaluation*, *68,* 100955. https://doi.org/10.1016/j.stueduc.2020.100955

Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean Journal of Anesthesiology*, *72*(6), 558–569. https://doi.org/10.4097/kja.19087

Lee, S., & Jones, P. (2019). Tutorial attendance and its relationship to exam performance: A quantitative analysis. *Journal of Educational Research, 40*(4), 511-525.

Lee, H., Chung, H. Q., Zhang, Y., Abedi, J., & Warschauer, M. (2020). The Effectiveness and Features of Formative Assessment in US K-12 Education: A Systematic Review. *Applied Measurement in Education, 33*(2), 124–140. https://doi.org/10.1080/08957347.2020.1732383

Martin, J., Adana, D. D. R. de, & Asuero, A. G. (2017). Fitting Models to Data: Residual Analysis, a Primer. In *Uncertainty Quantification and Model Calibration*. InTech. https://doi.org/10.5772/68049

O'brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity*, *41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

Salmerón Gómez, R., Rodríguez Sánchez, A., García, C. G., & García Pérez, J. (2020). The VIF and MSE in Raise Regression. *Mathematics*, *8*(4), 605. https://doi.org/10.3390/math8040605

Smith, T. (2017). The impact of tutorial participation on final exam scores: A cohort study. *Educational Studies Quarterly*, *25*(1), 67-82.

Weldmeskel, F. M., & Michael, D. J. (2016). The impact of formative assessment on self-regulating learning in university classrooms. *Tuning Journal for Higher Education*, *4*(1), 99. https://doi.org/10.18543/tjhe-4(1)-2016pp99-118

Wiliam, D. (2011). *Embedded Formative Assessment*. Solution Tree Press.