Global Higher Education Competitiveness: Analysis of Academic Reputation and Its Influencing Factors

Hexiang Li, Fuxian Xiao, Yuchen Lin, Tommy Cheng, Yang Sun

Section 1: Introductions(360 words):

Academic reputation has emerged as a critical criterion in the assessment of global universities. It influences institutional prestige, faculty recruitment, student enrollment, research initiatives, and global visibility (Hazelkorn, 2015). QS World University Rankings 2025 assigns each university an Academic Reputation Score (ARS) based on extensive academic surveys. The key factors influencing academic reputation remain unclear. Institutions have to grasp them in a competitive global education scene if they are to strategically place themselves.

This project examines which institutional features most strongly relate to academic reputation, using QS 2025 rankings data from Kaggle with over 1,500 universities and 10+ variables. For linear models, the academic reputation score is a continuous and ratio-scaled variable from 1 to 100. Citations per Faculty, Employer Reputation, Faculty-Student Ratio, Sustainability, Employment Outcomes, and Institution Size Classification include both continuous and categorical predictors. Before regression modelling, missing values and standardization of variables were addressed to ensure model suitability.

The study uses multiple linear regression to examine how institutional features influence academic reputation. By controlling for other variables, MLR allows us to estimate the marginal effect of each predictor, ensuring interpretability. The robustness of the model will be assessed with assumption checks and residual diagnostics. The objective is not only to fit a model but also to identify actionable levers for academic institutions by interpreting the structure of the regression coefficients. MLR is appropriate because the response variable is continuous and ratio-scaled, and the model provides interpretable coefficients.

However, MLR relies on several key assumptions: linearity, independence, normality, and constant variance of residuals. These will be checked with diagnostic plots to confirm MLR suitability.

This research extends prior studies on global rankings, research visibility, international collaboration, and global engagement are significant factors influencing academic reputation, as indicated by previous studies (Shin & Toutkoushian, 2011). Hazelkorn notes that universities increasingly adapt to rankings by focusing on research output and employability. (Hazelkorn, 2015). We thus hypothesize that Academic Reputation Scores are statistically correlated with several institutional factors, including research output, internationalization, faculty resources and employability, among other institutional indicators.

Unlike QS's subjective assessment system, our framework is evidence-based and replicable. Interpreting the coefficients can guide institutions aiming for global competitiveness.

Section 2: Data Descriptions (311 words):

We found this dataset on Kaggle and QS.com (Monfared, 2025). Furthermore, the QS World University Rankings 2025 were compiled through a data collection process that combined institutional submissions, large-scale global surveys.

According to QS Quacquarelli Symonds (2025a), universities submitted detailed student, staff, and institutional data through QS platforms like MovelN. QS surveyed over 151,000 institutions through academic and employer opinion surveys (QS Quacquarelli Symonds, 2025b). The usage of our original dataset is to help students, universities, and policymakers compare global institutions, which is closely related to our research direction.

Our response variable is Academic_Reputation_Score. This variable can be treated as a continuous and ratio scale variable, ranging from 0 to 100. It measures an institution’s academic standing in the international higher education landscape. Because this variable is continuous and ratio-scaled, Academic_Reputation_Score is suitable as a response variable in a linear model.

The histograms show that most variables are right-skewed, indicating many universities have lower scores in several areas. But the histogram of International_Research_Network is unique, and it is an even spread (See figures 1 and 2).

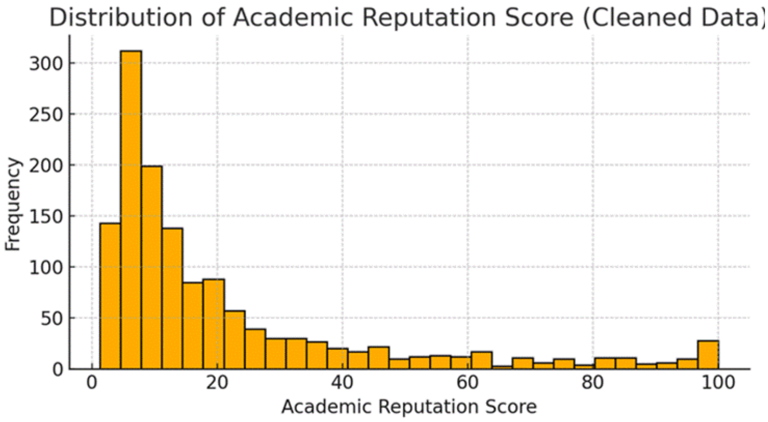

Figure 1. Histogram showing the distribution of Academic Reputation Score in the cleaned dataset. Data source: Manfared (2025).

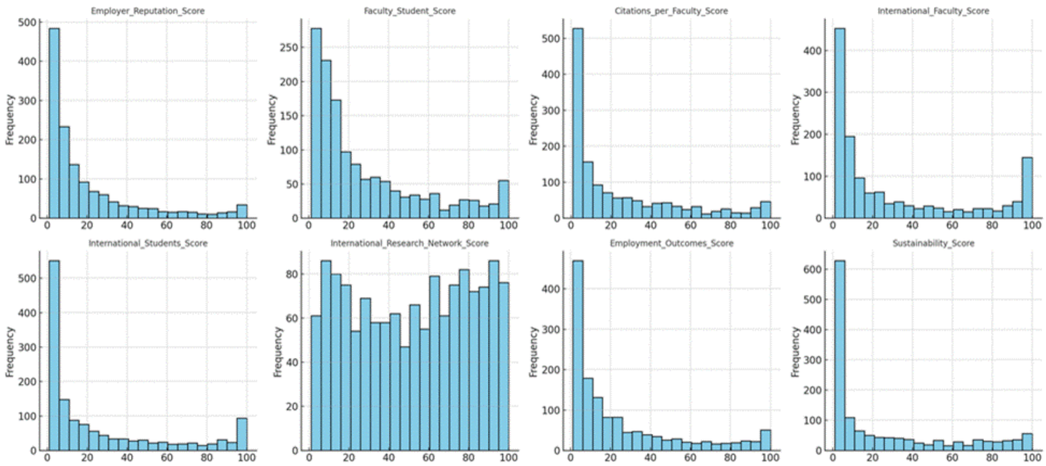

Figure 2. Histograms illustrating the distributions of the eight numerical predictors in the dataset. Data source: Manfared (2025).

Table 1 includes some descriptive statistics of 8 predictors and the only categorical variable, “SIZE”. Table 2 further details the categorical variable.

Numerical Variable Statistics

Variable	Mean	SD	Min	25th Quantile	75th Quantile	Max
Employer_Reputation_Score	20.892	24.423	1.100	4.500	27.700	100.00
Faculty_Student_Score	28.359	27.579	1.200	7.475	41.150	100.00
Citations_per_Faculty_Score	24.600	28.236	1.000	3.000	39.375	100.00
International_Faculty_Score	31.046	34.520	1.000	4.275	53.100	100.00
International_Students_Score	26.391	31.488	1.000	3.100	40.900	100.00
International_Research_Network_Score	51.641	29.605	1.000	24.500	78.200	100.00
Employment_Outcomes_Score	24.694	27.732	1.200	4.200	35.125	100.00
Sustainability_Score	25.913	31.646	1.000	1.500	44.600	100.00

Table 1. Descriptive statistics for the eight numerical predictors. Source: Monfared (2025).

SIZE Category	Original Count (n=1,376)	Cleaned Count (n=1,020)	Cleaned %
Small (S)	76	125	12.3%
Medium (M)	337	290	28.4%
Large (L)	639	455	44.6%
Extra-Large (XL)	324	150	14.7%

Table 2. Frequency and proportions of universities in each SIZE category (Excluding FO, FC, and CO because they are not directly comparable to the majority of universities in the dataset). Source: Monfared (2025).

Back to our research question: “Which factor is most influential on the response variable Academic_Reputation_Score?” After analyzing the histogram and statistical summary, International_Research_Network_Score is a strong candidate in our analysis. It measures a university’s level of international collaboration. Additionally, the mean of this variable, which is 51.641, is significantly higher than that of other variables, and its large variance suggests it has strong potential to explain differences in Academic_Reputation_Score.

Figure 3 indicates that many predictors, especially reputation scores, show linearity with Academic_Reputation_Score, supporting their selection.

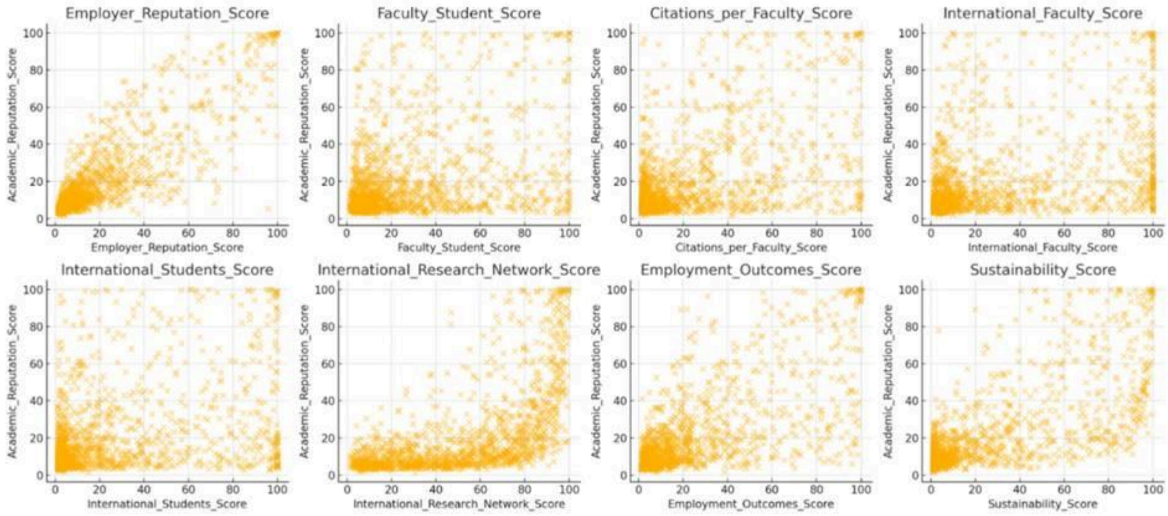

Figure 3. Scatterplots showing the relationship between each numerical predictor and Academic Reputation Score. Source: Manfared (2025).

In our cleaning process of the dataset, we removed 127 rows and 17 columns with missing values, and set 1,316 missing entries to 0.

<u>Section 3: Preliminary Results (312 words):</u>

We fitted a multiple linear regression with nine predictors. The model achieved an adjusted R-squared of 0.9465, a multiple R-squared of 0.9473 and a residual standard error of 10.16 on 953 DF. These values suggest a good fit in terms of explained variance and prediction accuracy. However, further diagnostics are required to check the assumptions of linear regressions.

Residual diagnostics reveal that the **Residuals vs Fitted values** plot shows a curvature in the red line and a funnel-shaped spread, suggesting violations of linearity and homoscedasticity. To identify suspicious contributing predictors, we examined the **Residuals vs Faculty_Student_Score** plot, which shows randomness and roughly constant spread, indicating compliance with assumptions. In contrast, the **Residuals vs SIZE** plot shows larger variability, indicating potential violation of homoscedasticity.

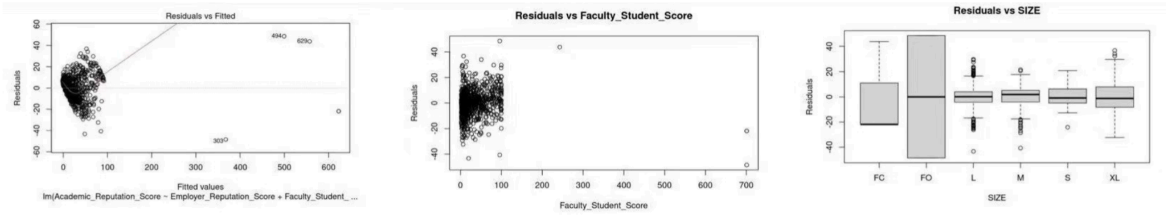

Figure 4. Residual analysis for the multiple linear regression model. (a) Residuals vs Fitted Values; (b) Residuals vs Faculty Student Score; (c) Residuals vs SIZE (categorical). These plots are used to assess linearity, homoscedasticity, and the influence of the categorical predictor SIZE. Data source: Manfared (2025).

For the normality assumption, the histogram and QQ-plot showed rough symmetry, showing a roughly symmetric shape with a heavier left tail, suggesting slight non-normality. For the independence of the residuals, we calculated the residual covariance matrix, which was not zero. Therefore, there might be some correlations between some of the residuals.

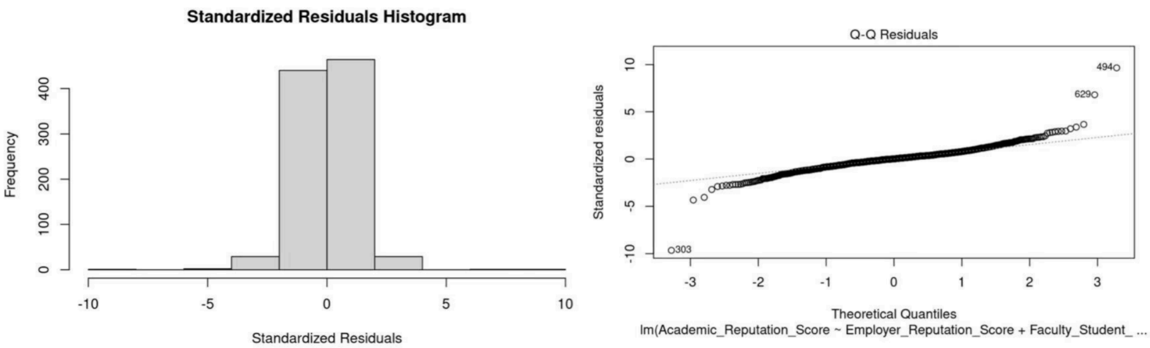

Figure 5: Residual analysis for the multiple linear regression model. a) Standardized Residual Histogram; b) QQ-plot of standardized residuals. Data source: Manfared (2025).

Next, we can apply the transformation on Academic_Reputation_Score to alleviate non-constant variance and non-normality. Also, we can try to drop any predictors to alleviate linearity issues. Moreover, time series analysis can be applied for independence.

The main conclusions in the literature about the factors influencing academic reputation are generally supported by the results of our early model. Visibility, international collaboration, and global engagement are significant factors influencing academic reputation. The idea is supported by these predictors ' significance and a high adjusted R-squared value of 0.9465.

Furthermore, Faculty_Student_Score has little explanatory power for Academic Reputation Score, although it meets the requirements of linear regression by displaying linearity and constant variance in residual analysis. This implies that indicators about teaching are mainly undetectable in evaluations of global reputation, despite being internally significant for the student experience.

Section 4: Model Selection (916 words):

4.1 Analysis process

First, in our first three sections, we use Academic_Reputation_Score as the response variable and the remaining nine variables as predictor variables to fit our preliminary model, **mlr_model**. As we mentioned earlier, mlr_model does not fit the data well, primarily because several predictor variables exhibit obvious right-skewed distributions, and some variables have p-values greater than 0.05. We used VIF to diagnose multicollinearity in the raw model.

	GVIF	Df	GVIF^(1/(2*Df))
Employer_Reputation_Score	8.697572	1	2.949165
Faculty_Student_Score	2.783861	1	1.668491
Citations_per_Faculty_Score	4.997183	1	2.235438
International_Faculty_Score	4.388435	1	2.094859
International_Students_Score	4.543237	1	2.131487
International_Research_Network_Score	5.867755	1	2.422345
Employment_Outcomes_Score	7.698585	1	2.774632
Sustainability_Score	6.550773	1	2.559448
SIZE	16.964246	5	1.327252

Table 3. The GVIF diagnostics for all pr edictors in the initial model.

The results show that the values of all predictor variables for GVIF^(1/(2*Df)) are less than the commonly used VIF threshold of 5, indicating that there are no serious multicollinearity issues in the model. However, to further enhance model simplicity and eliminate potential irrelevant predictors, we applied Lasso regression. After excluding ID-type variables (e.g., rankings and school names), we created a new model: **lasso_model**. This model retained all predictive variables from model_1 except for the International_Faculty_Score variable. Although Lasso only selected the dummy variable SIZEXL, we reintroduced the full SIZE categorical variable into the lasso_model to ensure interpretability, since categorical variables should be evaluated as a whole. The adjusted R^2 of lasso_model is 0.9466, indicating that this model has extremely strong fitting capability.

To further optimize the model structure following Lasso variable selection, we applied Stepwise Selection based on the AIC criterion, resulting in the updated **step_model**. To assess potential multicollinearity, we computed the adjusted GVIF values (i.e., GVIF^(1/(2*Df))) for all predictors. The results show that all adjusted GVIFs are below the commonly used threshold of 5, indicating that the step_model does not exhibit serious multicollinearity issues.

To further test how much improvement step_model has achieved in performance compared to the original mlr_model, we used the 10-fold cross-validation method to calculate the MSE of the two models. The results show that the MSE of step_model is 147.01, while the MSE of mlr_model is 197.11. The MSE of step_model is significantly lower than that of mlr_model.

At the same time, we conducted residual diagnostics. The Residuals vs Fitted plot (Figure 6a) shows a curved red line and a funnel-shaped pattern, suggesting violations of linearity and constant variance. The Scale-Location plot (Figure 6c) displays a slight upward trend, further indicating non-constant variance. The QQ-plot (Figure 6b) reveals minor deviations in the tails, suggesting potential non-normality. Additionally, the Residuals vs Leverage plot (Figure 6d) identifies a few high-leverage observations that may influence the model.

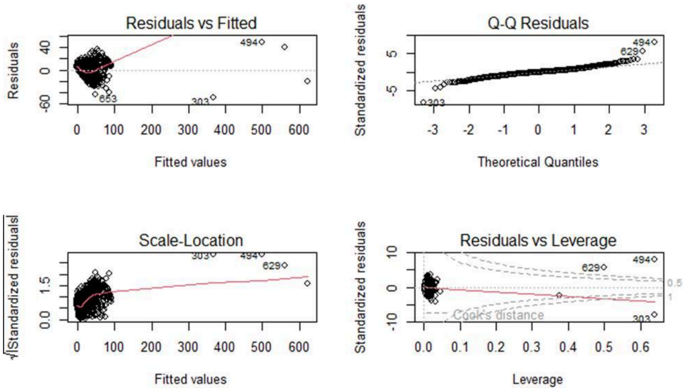

Figure 6. Diagnostic plots for the stepwise regression model: (a) Residuals vs Fitted, (b) Q-Q plot of residuals, (c) Scale-Location plot, and (d) Residuals vs Leverage.

These plots are used to assess model assumptions, detect non-linearity, evaluate normality and homoscedasticity of residuals, and identify potential influential observations.

	model 1	model 2	model 3	model 4
(Intercept)	4.380448	2.01979	3.95567	6.47730
Employer_Reputation_Score	0.514097	0.51435	0.51617	0.51617
Faculty_Student_Score	0.006258	0.00549	\\	\\
Citations_per_Faculty_Score	0.156358	0.15649	0.15432	0.15391
International_Faculty_Score	-0.011191	\\	\\	-0.01072
International_Students_Score	0.034830	0.02982	0.03000	0.03482
International_Research_Network _Score	0.071017	0.07143	0.07233	0.07205
Employment_Outcomes_Score	0.122381	0.12067	0.12081	0.12247
Sustainability_Score	0.061888	0.05926	0.06064	0.06334
SIZEFO	-84.41187	-82.81254	-83.91734	-85.59801
SIZEL	-7.130932	-4.88285	-6.75304	-9.15798
SIZEM	-8.309959	-6.06813	-7.85946	-10.24784
SIZES	-7.923003	-5.91966	-7.64797	-9.79938
SIZEXL	-2.741645	-0.43331	-2.37067	-4.84236
F statistics	\\	0.8468	0.5348	0.2881
p-value	\\	0.3577	0.5859	0.5915

Table 4. Comparison of regression coefficients, F statistics, and p-values across models 1 to 4.

To justify using step_model, we performed partial F-tests comparing it with models 2, 3, and 4. All p-values (0.3577, 0.5859, 0.5915) exceed 0.05, so we cannot reject the null hypothesis. This suggests that step_model performs similarly to more complex models, making it the preferred, simpler choice.

To further improve the homoscedasticity and normality of the model, we aim to apply the Box-Cox transformation method to the response variable Academic_Reputation_Score, and it suggests the best lambda for the response variable is 2/9 (0.2222222). Hence, we introduce our new model: **transformed_fit**.

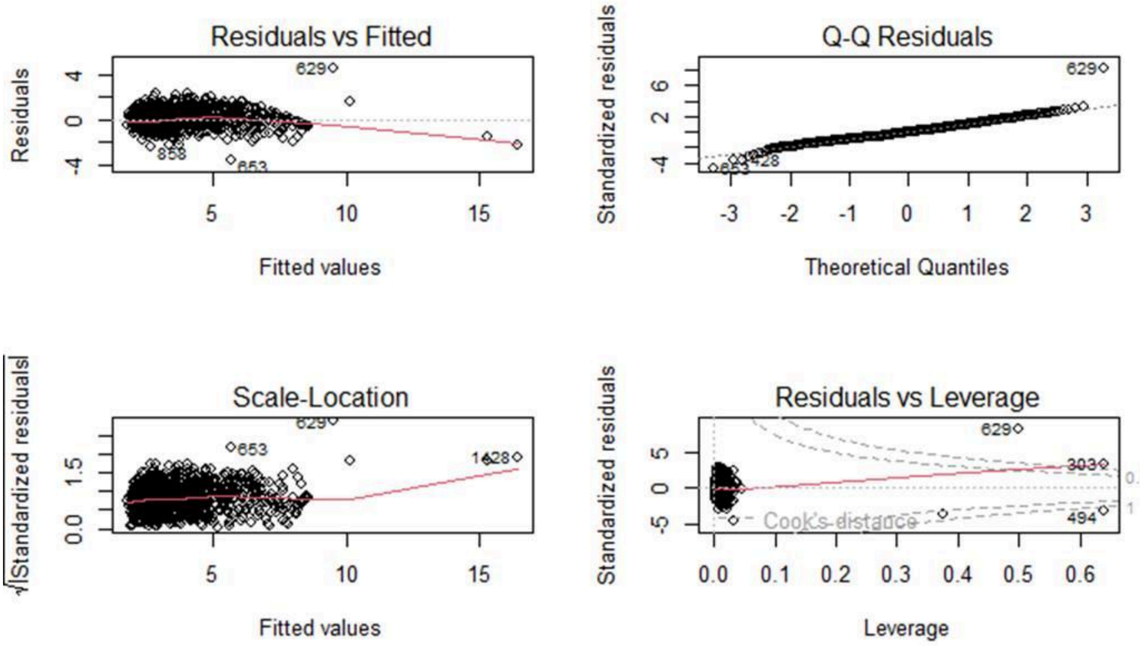

Figure 7. Model diagnostic plots for the final regression model: (a) Residuals vs Fitted values, (b) Q-Q plot of residuals, (c) Scale-Location plot, and (d) Residuals vs Leverage plot. These plots assess linearity, normality, homoscedasticity, and the presence of influential observations.

The diagnostic plots of the transformed model suggest that key assumptions are largely satisfied. Linearity and constant variance appear reasonable, with minor deviations due to a few outliers visible in the Residuals vs Fitted and Scale-Location plots (Figure 7a,7c). The QQ plot (Figure 7b) indicates approximate normality, while the Residuals vs Leverage plot (Figure 7d) reveals several high-leverage points that may impact prediction. Therefore, in the next step, we investigate outliers, influential points, and leverage to further improve model robustness.

We identified problematic observations using leverage values (64 high-leverage points), standardized residuals (4 outliers), Cook's Distance (multiple influential points exceeding 1.0), and DFFITS analysis.

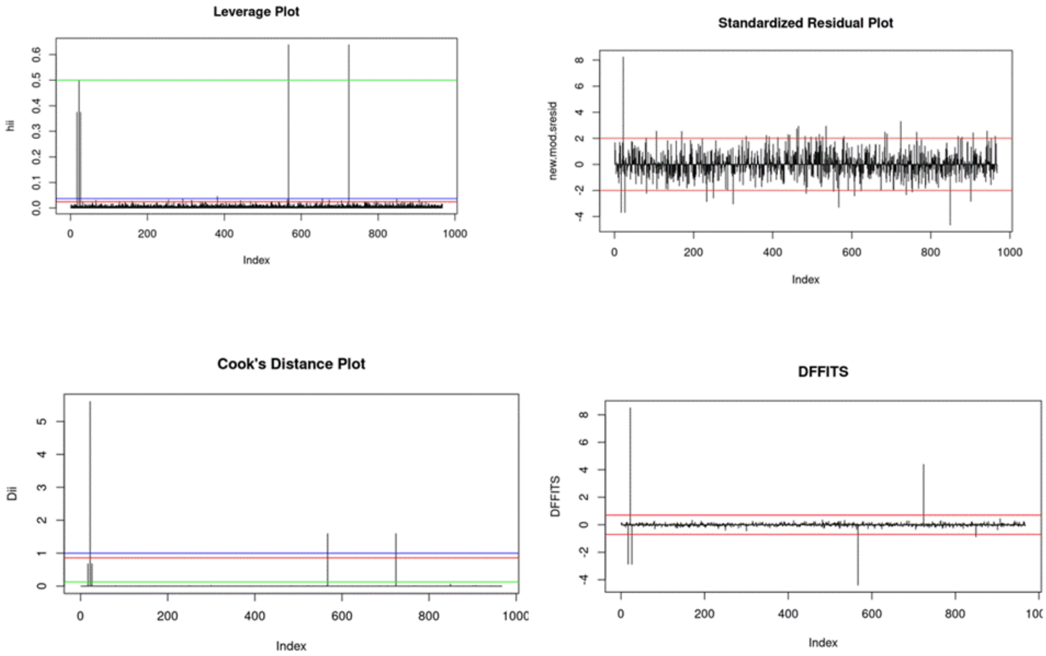

Figure 8. Influence and outlier diagnostics for the final regression model: (a) Leverage plot, (b) Standardized residual plot, (c) Cook's distance plot, and (d) DFFITS plot. These plots help identify high-leverage points, outliers, and observations with a strong influence on model estimates.

To evaluate the effect of high-impact observations, we removed ten data points with high Cook's Distance, DFFITS, or leverage values, resulting in the updated model **transformed_fit_cleaned**. Compared to transformed_fit, the cleaned model shows notably lower AIC, AICc, and BIC values, indicating improved model simplicity and fitting. Although its adjusted R^2 is slightly lower, the overall model fit and simplicity make transformed_fit_cleaned the preferred model (Table 5).

Model	adjusted R^2	AIC	AICc	BIC
transformed_fit (model 5)	0.8255187	2303.522	2303.903	2366.886
transformed_fit_cleaned (model 6)	0.8132442	2207.101	2207.380	2260.626

Table 5. Model comparison between the transformed model and the cleaned model.

To further validate the predictive capability of the final model, we compared the transformed_fit with the transformed_fit_cleaned using 10-fold cross-validation. The results showed that the MSE of the transformed_fit_cleaned was significantly lower than that of the original model (0.585 vs. 0.735), further confirming our conclusions.

In summary, we began with an initial model and systematically refined it using Lasso regression, stepwise selection, Box-Cox transformation, and influential points removal. At each step, we focused not only on goodness of fit but also on diagnosing issues such as non-constant variance, multicollinearity, and non-normality to ensure that the final model meets the basic assumptions of multiple linear regression. More importantly, after validation through cross-validation and information criteria, we found that the final model outperforms the initial model in terms of generalization ability and stability. Additionally, throughout the variable selection process, we integrated theoretical background and causal logic to avoid blindly relying on algorithmic variable removal, thereby enhancing the interpretability and credibility of the model's inferences. Therefore, we have good reason to believe that the optimized model not only better predicts the Academic Reputation Score but also more reliably reveals its primary influencing factors.

4.2 Our Final Model

$$
\begin{aligned}
\frac{9}{2}\left((\text{Academic\_Reputation\_Score})^{\frac{2}{9}} - 1\right) = {} & 1.6659459 - 0.0016129 \cdot I(\text{SIZE} = \text{SIZEM}) + 0.2842117 \cdot I(\text{SIZE} = \text{SIZES}) \\
& + 0.2158954 \cdot I(\text{SIZE} = \text{SIZEXL}) + 0.0351394 \cdot \text{Employer\_Reputation\_Score} \\
& + 0.0026094 \cdot \text{Citations\_per\_Faculty\_Score} - 0.0004698 \cdot \text{International\_Students\_Score} \\
& + 0.0139693 \cdot \text{International\_Research\_Network\_Score} + 0.0083028 \cdot \text{Employment\_Outcomes\_Score} \\
& + 0.0110356 \cdot \text{Sustainability\_Score}
\end{aligned}
$$

with the data from `train_cleaned`

<u>Section 5: Final model inference and results (998 words):</u>

5.1 Table summary for final model

To clearly present the results of multiple linear regression analysis, we have constructed a table that summarizes the key statistics for each predictor in our final model (See Table 6). The table includes the estimated regression coefficients, standard errors, t-values, p-values, and 95% confidence intervals.

Term	Estimate	Std. Error	t-value	p-value	95% CI Low	95% CI High
(Intercept)	1.6659	0.0668	24.9481	0.0000	1.5349	1.7970
SIZEM	-0.0016	0.0667	-0.0242	0.9807	-0.1326	0.1294
SIZES	0.2842	0.1203	2.3631	0.0183	0.0482	0.5202
SIZEXL	0.2159	0.0647	3.3367	0.0009	0.0889	0.3429
Employer_Reputation_Score	0.0351	0.0015	23.4761	0.0000	0.0322	0.0381
Citations_per_Faculty_Score	0.0026	0.0011	2.2986	0.0217	0.0004	0.0048
International_Students_Score	-0.0005	0.0010	-0.4838	0.6287	-0.0024	0.0014
International_Research_Network_Score	0.0140	0.0014	10.2682	0.0000	0.0113	0.0166
Employment_Outcomes_Score	0.0083	0.0012	6.7343	0.0000	0.0059	0.0107
Sustainability_Score	0.0110	0.0013	8.7142	0.0000	0.0086	0.0135

Table 6. Estimated coefficients, standard errors, t-values, p-values, and 95% confidence intervals for each predictor in the final multiple linear regression model.

The table reveals that several predictors have statistically significant relationships with the response variable (Academic_Reputation_Score), as indicated by a p-value less than 0.05. Specifically, after our model-selection procedure, SIZES, SIZEXL, Employer Reputation Score, Citations per Faculty Score, International Research Network Score, Employment OutcomesScore, and Sustainability Score are all significant predictors. Among these, Employer Reputation Score has the largest t-value, suggesting it is a particularly strong predictor. The 95% confidence intervals for these significant variables do not cross zero, further supporting their influence.

Conversely, SIZEM and International Student Score have high p-values and confidence intervals that include 0, indicating that their effects are not statistically significant in the presence of other predictors. This suggests that, after accounting for the other factors in the model, these variables do not have a meaningful impact on the response variable.

The use of confidence intervals alongside p-values offers additional context, showing the likely range of each coefficient's true value in the population. Standard errors reflect the precision of the estimates—smaller standard errors indicate more reliable coefficients.

By including all these in a single table, we provided a transparent and comprehensive summary of the final model. This table supports both statistical interpretation and practical decision-making, allowing readers to see at a glance which factors matter most and how certain we are about each result.

5.2 Interpretation of Model Coefficients

Each coefficient in our final multiple linear regression model represents the expected change in the transformed Academic_Reputation_Score for a one-unit increase in the predictor, holding all other variables constant. Interpreting these coefficients in context helps us understand how different institutional factors contribute to a university's academic reputation.

1. **Employer_Reputation_Score** (estimate: 0.0351, $p < 0.001$): This predictor is the strongest, with a high t-value as well as a very small p-value. For each additional point in employer reputation, the transformed academic reputation score increases by 0.0351, holding all else equal. This suggests that employer perceptions are highly influential in shaping a university's academic reputation.

2. **SIZES** and **SIZEXL** (0.2842 and 0.2159, respectively): These two predictors are also two strong candidates for our research question since they have a very large coefficient. In context, large and extra-large institutions have higher academic reputation scores compared to the baseline size category. This could indicate that larger institutions are more visible and better resourced, positively affecting their reputation.

3. **Citations_per_Faculty_Score** (0.0026, $p = 0.0217$): Each additional point in the citations per faculty score is associated with a small but significant increase in academic reputation, reflecting the importance of research impact.

4. **International_Research_Network_Score** (0.0140, $p < 0.001$): Universities with stronger international research networks tend to have higher academic reputation scores, indicating the value of global research collaboration.

5. **Employment_Outcomes_Score** (0.0083, $p < 0.001$): Higher employment outcomes for graduates are linked with increased academic reputation, possibly because

successful alumni enhance the perceived quality of the university.

6. **Sustainability_Score** (0.0110, $p < 0.001$): Universities with greater emphasis on sustainability also tend to score higher in academic reputation, which may reflect growing awareness of social responsibility in academics.

7. **SIZEM** and **International_Students_Score** were not statistically significant ($p > 0.05$). And the coefficients are negative, indicating that each additional point in these two items is associated with a small decrease in academic reputation. This suggests that after accounting for other factors, medium size and the proportion of international students do not significantly impact the academic reputation in this dataset.

Our analysis reveals that "Employer Reputation Score, "SIZES", and "SIZEXL" are three candidates for the final research question. Therefore, in practical terms, this suggests that universities seeking to improve their academic reputation should not only focus on growing their size or research output, but also prioritize building meaningful employer partnerships and enhancing the career readiness of their graduates. For smaller institutions, building strong relationships with employers can help make up for not being a large organization. For larger institutions, investing in both research excellence and workforce relevance creates a competitive advantage.

5.3 Assessment of Model Performance

To evaluate the performance of our final multiple linear regression model, we considered several commonly used metrics: R-squared, Adjusted R-squared, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). Each of these metrics offers different insights into how well our model fits the data and its overall quality. Below is the table of the final model performance metrics:

Metric	Value
R-squared	0.8150
Adjusted R-squared	0.8132
AIC	2207.1012
BIC	2260.6260

Table 7. Performance metrics for the final model, including R-squared, Adjusted R-squared, AIC, and BIC, which summarize the model's fit and complexity.

R-squared for the final model is 0.8150, indicating that about 81.5% of the variance in the transformed Academic Reputation Score is explained by the predictors in the model (Table 7). This is a strong result for a model with multiple predictors and suggests a good fit to the data. However, R-squared alone can be misleading, as it always increases with the addition of more variables, regardless of whether those variables are meaningful.

To address this, we consider Adjusted R-squared (0.8132), which adjusts for the number of predictors relative to the number of observations. The small difference between R-squared and Adjusted R-squared suggests that the predictors included in the final model are all contributing meaningful information, and the model is not overly complex or overfitted.

AIC (2207.10) and BIC (2260.63) are both information criteria that penalize model complexity. Lower value indicates better model quality, balancing goodness-of-fit with model simplicity. In our analysis, we compared models with and without influential points, as well as models with different predictor sets. The final model achieves relatively low AIC and BIC values compared to more complex alternatives, indicating a favourable balance between fit and simplicity.

In summary, high R-squared values show our model explains much of the variance in academic reputation, while low AIC and BIC indicate good fit without excess complexity. Together, these metrics support the model's validity for interpretation and prediction.

Section 6: Discussion and conclusion (985 words):

The aim of this project was to find the most important factors that contribute towards a university's Academic Reputation Score in the QS global ranking dataset. We improved our multiple linear regression model by employing a systematic modelling process consisting of variable selection using Lasso and stepwise AIC, Box-Cox transformation of the response variable, and identification and elimination of influential and high-leverage points. This improved model interpretability and assumption alignment. The final model accounts for roughly 81% of the variation in the adjusted Academic Reputation Score, underpinned by an Adjusted R-squared of 0.8132, supporting the model's ability to explain reputation determinants (Sheather, 2009; Faraway, 2002).

6.1 Key Findings in Context

Our findings confirm that Employer Reputation Score is the most influential factor driving academic reputation, with a strong positive coefficient and a highly significant p-value, emphasizing employer perceptions. Perceived employability emerged as the most important factor, validating QS's emphasis on market outcomes(Marginson, 2006).

The International Research Network Score was another strong predictor. These institutions tend to be more visible globally, and research productivity, as shown by Citations per Faculty, also contributes significantly to reputation. However, the International Students Score and Faculty Student Ratio did not significantly affect academic reputation despite its role in QS rankings. These reinforce earlier critiques that ranking inputs may lack explanatory value and distort performance emphasis (Dill & Soo, 2005; Docampo, 2012).

Institution size was also significantly (after transforming the response variable with the Box-Cox method) predictive of academic reputation. Specifically, Extra-large institutions (SIZEXL) were associated with higher reputation scores ($p < 0.001$), possibly due to greater visibility and resources. In contrast, still neither medium-sized (SIZEM) nor small-sized (SIZES) institutions showed significant differences from the baseline category (likely large size) in terms of transformed academic reputation scores ($p = 0.794$ and $p = 0.094$, respectively). This nuanced result suggests that the influence of size on reputation may be nonlinear. While student diversity and institution size may intuitively seem important, they do not exert strong effects on academic reputation when other variables are held constant.

Finally, both the Sustainability Score and the Employment Outcomes Score were statistically significant. This indicates that modern academic reputation extends beyond traditional metrics like citations or faculty ratios. Universities seen as socially responsible and successful in graduate employment tend to enjoy stronger reputations. As global problems like climate change and job uncertainty continue to influence public narratives, the reputational advantage may favour those institutions perceived as social actors (Hazelkorn, 2015).

6.2 Implications and Strategic Recommendations

These predictors suggest that institutional reputation is no longer solely based on traditional academic metrics, but increasingly on a broader ecosystem of stakeholder perceptions, global engagement, and research impact. This supports Hazelkorn's (2015) call for data-informed strategy at both institutional and policy levels.

Universities should enhance visibility and graduate outcomes through employer partnerships such as internships and advisory boards. Similarly, international research collaboration, as reflected in the International Research Network Score, was a positive predictor, reflecting that research flows across national borders are central to reputational capital. Institutions are encouraged to expand faculty exchange and co-authored publications. Such interventions, as Marginson (2006) contends, can boost cross-border citation rates and scholarly presence.

Sustainability and employment outcomes show that societal relevance boosts reputation. Schools that are environmentally conscious and effective in preparing students for the workforce are increasingly progressive, highlighting their contributions beyond academic metrics (Hazelkorn, 2015). These findings support prior critiques that not all traditional QS inputs—such as faculty ratios or international student counts—are valid predictors of academic reputation."

Finally, universities and other groups should continue to be critical of ranking systems. Not all QS ranking metrics are supported as valid predictors of reputation. Instead of blindly adopting external benchmarks, it might be better to create internal performance indicators that are based on facts and are in line with the institution's purpose.

6.3 Limitations and Future Research

Although our final model exhibits robust performance and interpretability, there are still numerous constraints. Despite the Box-Cox transformation and outlier removal, mild deviations from homoscedasticity and normality remain, particularly among high-leverage points. This suggests that linear regression may not completely capture the non-linear patterns or interaction effects that are inherent in the data.

There was a serious limitation to the range of data available from the QS agency. Qualitative feedback from students or faculty, research funding, academic awards, and the quality of teaching/learning could have been included. Though hard to quantify, these can greatly affect institutional standing. Hazelkorn (2015) and Rauhvargers (2011) noted a potential problem in the use of quantifiable metrics if it masks the qualitative dimensions of an institution's performance.

Moreover, our model was confined to the assumption that it assumes fixed and additive effects. Future work could use interaction terms or regularization (e.g., Ridge, Lasso) to explore variable interplay (James et al., 2021).

Finally, our evaluation used a cross-sectional method. There is a potential influence of leaders, policies, and external disruptions upon a system's reputation, and longitudinal data could enhance our understanding of how those influences wax or wane over time.

6.4 Conclusion and Final Reflection

In conclusion, this project has produced a statistically robust and interpretable model. A series of diagnostics were carried out to ensure that the principal key assumptions of linear regression, namely linearity, normality and homoscedasticity, were satisfied, which resulted in a credible analysis and replicable findings. Residual diagnostics found a good fit and reliability of the final model.

Substantively, we show that reputation reflects not just research or size, but global engagement and societal relevance. Employer interaction, global research, and societal contribution are essential elements in forming a global academic reputation. These findings resonate within broader literature, which argues that institutional reputation is increasingly reliant on external relationships, social responsibility and global connection (Hazelkorn, 2015; Marginson, 2006). Our model not only makes an empirical contribution but also provides organizations with strategic direction to inform performance-based evaluations in a changing societal context.

Section 7: Author Contributions:

Yuchen Lin: Responsible for data cleaning, constructing the MLR model, and formatting the final proposal. Completed Section 4 writing.

Hexiang Li: Performed residual analysis and summarized key findings. Provided code for model selection and inference. Assisted group members in refining their written reports. Made corrections for Sections 1-3. Completed Section 4 writing.

Tommy Cheng: Completed partial revisions in Sections 1–3.

Yang Sun: Completed Section 2 and Section 5 writings. Made partial corrections for Sections 1-3 based on the feedback of the Proposal. Responsible for the format consistency (e.g., in-text references and captions for figures/tables, font) in the final project report.

Fuxian Xiao: Responsible for writing Section 6 (Discussion and Conclusion), comparing the model to relevant literature, compiling and formatting references, and participating in the final revision.

Section 8: References:

Dill, D. D., & Soo, M. (2005). *Academic Quality, league tables, and public policy: A cross-national analysis of University Ranking Systems.* Higher Education. https://doi.org/10.1007/s10734-004-1746-8

Docampo, D. (2012, July 1). *Reproducibility of the shanghai academic ranking of world universities results - scientometrics*. SpringerLink. https://doi.org/10.1007/s11192-012-0801-y

Faraway, J. J. (2002). *Practical Regression and Anova using R.* University of Bath.

Hazelkorn, E. (2015). *Rankings and the Reshaping of Higher Education*. https://doi.org/10.1057/9781137446671

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2022). *An introduction to statistical learning: With applications in R*. Springer.

Marginson, S. *Dynamics of National and Global Competition in Higher Education. High Educ 52,* 1–39 (2006). https://doi.org/10.1007/s10734-004-7649-x

Monfared, M. (2025). *QS World University Rankings 2025* [Data set]. Kaggle. https://www.kaggle.com/datasets/melissamonfared/qs-world-university-rankings-2025

QS Quacquarelli Symonds. (2025a). *Data collection process for QS World University Rankings 2025* [Institutional Data Submissions]. https://www.qs.com/data-collection-qs-rankings

QS Quacquarelli Symonds. (2025b). *Data collection process for QS World University Rankings 2025* [Academic and Employer Surveys]. https://support.qs.com/hc/en-gb/articles/4405955370898-QS-World-University-Rankings

Rauhvargers, A. (2011). *Global university rankings and their impact.* European University Association

Sheather, S. (2009). *A modern approach to regression with R*. SpringerLink. https://link.springer.com/book/10.1007/978-0-387-09608-7

Shin, J. C., & Toutkoushian, R. K. (2011b). The past, present, and future of University Rankings. *University Rankings*, 1–16. https://doi.org/10.1007/978-94-007-1116-7_1