A Statistical Metric of University Quality Based Upon the NCES-IPEDS Data

William D. Wachob

William.Wachob@mi.rr.com

Lawrence Technological University

Southfield, Michigan 48075, USA

Abstract

The development of a quality metric for colleges and universities is an area that causes a great deal of conversation each year. Metrics used by most colleges and universities are normally localized to the school. This is done in order to make them relevant to the desired user or outcomes. When evaluating the quality of a college or university, there are a number of potential factors to be considered. This makes the whole process unwieldy. This paper will focus upon the development of an academic quality metric based upon the national standardized tests. If a simple single metric could be identified, many other comparisons can then be made. This paper develops the concept of an rScore that is a ranking score based upon the standardized testing scores of the incoming freshman class. The standardized testing scores are self-reported by the institutions in their NCES-IPEDS (National Center for Educational Statistics - Integrated Postsecondary Education Data Set) submissions.

Keywords: quality, ranking, metric, IPEDS, rScore

1. INTRODUCTION

There is an initial premise in this work which is represented in the following figure:
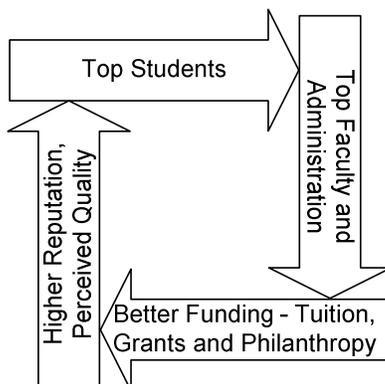

Figure 1 - College Funding "Arms Race"

When we read the arrows as "attracts", the intent of the overall process becomes clear. The relationship between these components is described in a paper by Cunningham and Cochi-Ficano (2002). This process also represents the basic funding philosophy of many college and universities. It results in a fundamental belief that causes extra focus to be placed on grants, philanthropy and tuition revenue. The "war chest" of endowed funds for some colleges is very large. This allows that university far more freedom in enhancing their faculty, facilities and image. With sufficient application of financial resources, institutions can effectively outspend their competitors. That spending can be used to build new or remodeled facilities and update laboratory and computer equipment. It can even be used to subsidize proposal efforts for additional project funding. The spending can also be used to incent top faculty and researchers to come to or stay with a school. This practice is the same when we look at private industry. While private industry does not generally have endowed assets, the corporate coffers most likely will have a "war chest" that can be used to make similar strategic moves when they are deemed necessary. This concept is discussed in Fuller (1993).

The process described in Figure 1 relies heavily upon the element of higher reputation or perceived quality. As there is

no overall specific metric for rating the quality of colleges and universities, a statistical methodology for depicting the quality of an institution would be useful. One primary area where such a metric could be used is in institutional reporting and analysis. The metric would provide another means to describe competitive institutions. If one school wished to emulate a particular success of another, one element of their discussion would likely be an assessment of whether the students at one school or the other are likely to have better success in that particular discipline.

The concept of perceived quality carries through to the needs of the public and perspective students. Several college selection publications (see references) advise the student to include "quality" in their college selection process. This will largely default to the unscientific lists of "best" colleges and marketing from the school. This project proposes a common metric that could be used in the decision process of college selection.

2. FOCAL THEORY

The focal theory of this work is that it is possible to express a measure of the quality in a single numeric score. This score has been named the "ranking score" or rScore for short. The institutions can also be grouped into bins for the purpose of analysis and comparison. These are called the "ranking bins", or "rBins". Both of these terms can be shaped to bias them towards the type of institutions you wish to favor. For example, developing a metric that gives higher weight to engineering / scientific schools is basically a matter of giving higher weights to the math scores of the SAT (Scholastic Aptitude Test) or the ACT (originally American College Test, changed to simply ACT in 1996). One definition of the rScore is as follows:

$$\frac{(2*\mathrm{SAT_{Math25}}+\mathrm{SAT_{Verbal25}})*\mathrm{ACT_{Composite25}}}{100}$$

Equation 1 – rScore25 calculation

The SAT and ACT scores were obtained as data reported by the colleges for their incoming freshman class. The data only includes students who are attending college for the first time. The choice of the elements was made to create a small bias towards the SAT mathematics score indicating that the students were likely to be more oriented towards science and engineering. The choice of the 25th percentile was made as that would indicate that 75% of the students had achieved the reported score or better. This effectively provides a measure of the likelihood of success of the incoming freshman class in mathematics and science.

It should be noted that the actual rScore for an institution is less important than the use of a repeatable methodology with which to rank or group the institutions under analysis.

Creating bins for the rScore values also suggests that the data set could be broken into a limited number of groups. Each group can then be shown to be statistically different from the others using the Student's t-statistic as a test indicator.

A more complete discussion of the Student's t-statistic is provided in two references (Anderson (2002) and Lipson (1973)). The equation for calculating the t-statistic is:

$$t - statistic \; = \frac{abs\left(\bar{x}_1 - \bar{x}_2\right)}{\sqrt{\dfrac{\sigma_1^{\,2} + \sigma_2^{\,2}}{n}}}$$

Equation 2 - Student's t-statistic

When n is sufficiently large the t-statistic values of interest become:

% Probability	t-value
95%	1.96
98%	2.33
99%	2.58

When the t-statistic is calculated, a score that exceeds the above values suggests that the two sets come from different populations.

The methodology followed in this work is modeled on the methodology used in Tang (2004). In this work, the authors compared a "university quality" metric to the compensation of the university CEOs, tuition and operating expenditures. They were able

to develop equations of these metrics and relate them to perceived university quality metrics. This work used the reputation rankings as published each year by the Newsweek organization. Tang (2004) expands upon the use of such a metric. They key relevance of this work is that it establishes that there is a statistical relationship between certain "perception metrics" and hard data associated with the colleges and universities they studied.

3. HYPOTHESIS

The hypotheses associated with this analysis were constructed such that they could support the development of the focal theory. These hypotheses were that:

- A relationship exhibiting a high degree of correlation exists between the rScore and admission rates of the institution [AdmissRate=f(rScore)]
- A relationship exhibiting a high degree of correlation exists between the rScore and faculty salaries [FacSalaries=f(rScore)]
- A relationship exhibiting a high degree of correlation exists between the rScore and the tuition charged by the institution [Tuition=f(rScore)]
- A relationship exhibiting a high degree of correlation exists between the rScore and the endowment assets of the institution [Endowment=f(rScore)]

A high degree of correlation would be accepted when the correlation coefficient was above 0.6. Earlier work by Wachob (2009) identified that the variability in reporting some the data values precluded using these forms for predictive modeling. This does not mean that the existence of a relationship cannot be shown.

4. DATA ACQUISITION

To obtain the data for the basic calculations, it was necessary to access the National Center for Education Statistics – Integrated Postsecondary Education Data Set (referred to as simply IPEDS). The data year used was 2005-2006 as that was the latest year that had a full set of data. The data from the Dataset Cutting Tool (DCT) was used to extract all of the school information. The data fields extracted from the data set are shown in attachment 1. The database was

then divided into several related tables as shown in the following table:

Data Table	Contents
inst	Basic information about the institution
endow	Value of the endowments of the institution
enroll	Enrollment counts, number of applications, number of acceptances and enrollments
salary	Number of faculty employees, average salary
staff	Number of staff employees and total salary
test	Reported scores for SAT and ACT standard tests
tuition	Reported tuition, fees and housing costs

Table 1 - MySQL data tables

The data tables were unified using the IPEDS identification number for the institution, denormalizing the data set in the interest of convenience of use.

The initial data set consisted of 7,018 institutions. A number of filters were then applied to the data set to limit the number of institutions that would be considered. These filters were:

- Only four-year degree-granting institutions were included
- Only institutions that reported endowment figures for 2005-2006 were included
- Only institutions that reported their 25th percentile SAT and ACT scores were included
- Only institutions that reported faculty salaries for 2005-2006 were included
  Application of these filters yielded a set of 1,063 schools. These schools had the following degree characteristics:

Degree Type	Number
Bachelor's	474
Master's	389
Doctoral	200

Table 2 - Distribution of data set by degree

Examination of the rScore25 values shows the following distribution:

rScore (min)	rScore (max)	Number
0	100	4
100	200	100
200	300	552
300	400	275
400	500	83
500	600	44
600+		5

Table 3 - Distribution of institutions by rScore25

The shape of this distribution is what would be expected. There are fewer schools that reflect the higher rScores. A further analysis shows that only 12.4% of the institutions had a score of 400 or more.

In order to make data handling more manageable, as further subset of the 1,063 schools was chosen. A nominal sample size of thirty-five (35) institutions was chosen. This choice results in selection of 3.3% of the schools. The selection of the thirty-five schools was made using the following algorithm:
1. Partition the data set into groups by the type of degree being granted
2. Rank order the resulting set in descending order by rScore25 values
3. Calculate the number of samples for that group by taking 3.3% of the group size
4. Round the number of samples to an integer
5. Beginning with the highest ranked sample, select every rth sample where r was the value calculated in step 4.

When the process described above is followed, the number of samples becomes:

Classification	Number of Schools	Number of Samples
Bachelor's	474	16
Master's	389	13
Doctoral	200	7

Table 4 - Sample sizes by degree type

A fourteenth sample was added to the Master's classification in order to include the author's institution. This accounts for the 37th sample. A full listing of the sample set is included in appendix I. The identification of the source institutions has been omitted.

5. DATA ANALYSIS

To make the data analysis, the segments represented in figure 1 need to be mapped into available datasets. The datasets available through the National Center for Educational Statistics – Integrated Postsecondary Education Data set provided the necessary information.

The data from the thirty-seven samples was processed for linear correlation using Minitab and Excel with the following results:

	Slope	Intercept	r^2	r
Admiss Rate	-0.1115	103.3	0.473	0.688
Faculty Salary	120.4	2,0183.0	0.646	0.804
Tuition	61.1	5,154.9	0.468	0.684
Endowment	3.354e6	-849.9e6	0.659	0.812
Log_{10}(Endow)	0.00546	5.914	0.564	0.751

Table 5 - Linear correlation results

From these results we can see that the hypothesis of a strong correlation between the rScore25 and the reported observations provided by the institutions holds true. The data and trend lines for each set were also graphed and are included in the appendices as figures 2 through 6. The data for the endowed assets was transformed by taking the log10 of the reported endowed assets before plotting.

Admissions Rate

Four-year degree granting institutions typically have enrollment policies that are based in part upon the likelihood of success for the student. Components of that assessment often include metrics such as the student's high school grade point average and performance on one or both of the standardized national tests (ACT and SAT). Schools that have better reputations for their academic programs have the ability to be more selective in whom they admit to their programs. When the admissions rates are compared to the rScore statistic a good correlation factor of 0.688 exists. This is shown in Figure 2 - Admissions Rate. The negative value of the slope is consistent with the observation that admissions rates fall as the rScore increases. If the rScore is a reflection of the likelihood of success for the students, then the idea that schools with higher rScore values can be more selective

in their admissions policies is supported. Among the selected schools, institutions with "open enrollment" policies were not included. This is not a major limitation because relatively few four-year degree granting institutions have open enrollment practices.

Faculty Salaries

The indicator used for the analysis was the reported faculty salaries adjusted for a nine-month school year. In this case, we find that the positive slope of the relationship is an indicator that the reported salaries increase with increasing rScores. This data is represented graphically in Figure 3 - Faculty Salaries.

While higher salaries do not necessarily mean that instruction is better, it is also likely that better known professors will draw the students with higher SAT and ACT scores. The reasoning here becomes circular except for the observation that the independent axis of this relationship is determined by the incoming freshman class. The rScore for the class is based upon standardized scores that are often made in the high school junior year.

The higher correlation between the rScore and the faculty salaries (0.804) may also be influenced by the higher tuition costs for more popular schools.

Tuition

Although the relationship between rScore and tuition has the lowest correlation coefficient, it still exhibits a good affinity based upon its r value of 0.684. The tuition values were not adjusted for geographic differences across the country. This data is shown graphically in Figure 4. The indications also do not represent any adjustments for market related factors. The reported values are made up of two components. Tuition, books and fees that are relatively constant within a region comprise the initial element. Local housing costs represent the other element. These have a higher degree of variance across geographic regions.

Endowed Assets

An initial analysis of the endowed asset data was normalized into a per student relationship. The statistical analysis failed to support a high correlation coefficient. A second analysis was performed using only the raw reported endowment amounts. This relationship had the strongest correlation of the data analyzed. When the data was plotted using the log10 of the reported amounts the relationship also had good correlation. This is depicted graphically in **Error! Reference source not found.**. This graph also suggests that a higher order relationship may exist between the rScore and tuition data. Fitting the values into a quadratic relationship (Figure–6) showed an exceptionally high correlation (0.902).

6. CONCLUSIONS

The initial premise of this paper was that a statistical metric could be constructed to indicate the quality of an institution. Webster's dictionary defines quality as a "degree of excellence" or "superiority of kind". In assessing higher educational institutions, many factors come to bear on the definition. It is likely that top students, faculty and benefactors will affiliate with institutions that have a better reputation. While an important element of this is the brand management of the institution, the never ending relationship depicted in Figure 1 still holds true. When the data is analyzed, a strong statistical relationship can be shown between the rScore25 and appropriate observations to the segment.

One of the observations from the rScore is that the distribution of the results follows a normal distribution. The histogram of the distribution of the rScores of the sample data set is provided as figure 7.

The rScore statistic is perhaps most useful when performing comparison analysis between institutions. Small differences (<5%) should be considered as noise in the analytical process. Large differences raise the question of capability between the comparison incoming freshman classes. In this regard grouping of the institutions into rBins becomes an appropriate tool.

The listing of the top 10 schools as calculated by the rScore is included in Appendix VII. The list represents the rScore as calculated with the formula in equation 1. This version of the rScore gives higher weighting to the SAT math25 score and

results in a list that has a higher number of technology schools. This does not rule out liberal arts programs as many students who score well in the math component also score well in the verbal and the ACT composite25 scores.

The rScore methodology provides a quick tool for evaluating the academic strength of an institution based upon the likelihood of success for their freshman class. The "arms race" hypothesis depicted in figure 1 has been validated against the rScore with a good level of correlation. The rScore can then be used in modeling to investigate its relationships to other common metrics. For example, the rScore has been correlated with technology infrastructure spending. The results were not as consistent as would be necessary for empirical modeling, but medium sized subsets of the results have shown promise.

Beyond using a statistic such as this for college selection decision making, the question becomes what, if anything a school could do to improve their score? The data from this study suggest that a significant increase in endowments would be the best place to focus. This is hardly surprising as it is one of the critical elements in driving the whole "arms race" cycle.

The component of faculty salaries represents the quality of the faculty. Recruitment of top faculty is normally an objective for any school. The need to maintain salary structures and competitive tuition pricing will act as a dampening mechanism in this element of the process.

The acceptance policies of the school also play a part in this cycle. By increasing the SAT and ACT scores required for admission, the rScore could be increased. Student tuition has a similar relationship in this process.

Increases in tuition and tightening of acceptance policies would undoubtedly have a negative effect on enrollment numbers. Given the importance of having a critical mass of students to sustain school operations, this may be counter-productive.

7. REFERENCES

Anderson, Sweeney, Williams (2002), *Statistics for Business and Economics,* South-Western division of Thompson Learning, pp 560-567

Carter, Carol C. (2006). College Selection Criteria: Reading Between the Lines. Retrieved June 2006, from http://www.lifebound.com/lifebound-resources-for-parents/college-selection-criteria

Coates, Joseph F., Mahaffie, John B., Hines, Andy (1997), 2025: Scenarios of US and Global Society Reshaped by Science and Technology, Oakhill Press, Greensboro, NC, ISBN 1-886939-09-8

College Selection Criteria (2006), Retrieved June 2006, from http://bcsd.k12.ny.us/high/counseling/collselection.htm

Cunningham, Brendan M., Cochi-Ficano, Carlena K. (2002), The Determinants of Donative Revenue Flows from Alumni of Higher Education, The Journal of Human Resources, Vol. XXXVII, No. 3, pp 540-569

DePauw University Admissions (2006), College Selection Criteria, Retrieved June 2006, from http://www.depau.edu/admission/applying/tips/college-search/criteria.asp

Fuller, Mark B. (1993), Business as War, *Fast Company Magazine*, October 1993, Issue 00, pp 42, Retrieved December 2006 from http://www.fastcompany.com/online/00/war_printer_friendly.html

Highland Park High School College Planning Handbook (2006), Criteria for College Selection, Retrieved June 2006, from http://www.dist113.org/hphs/handbook/bcriteria.htm

Libby, Wendy (2005), College choice is more than a numbers game, Columbia Daily Tribune, Published on May 1, 2006, Retrieved June 2006 from http://www.showmenews.com/2005/May/20050501Comm008.asp

Lipson, Charles and Sheth, Narendra J. (1973), *Statistical Design and Analysis of Engineering Experiments,* McGraw-Hill Publishing

Pascarella, Ernest T. (2001), Identifying Excellence in Undergraduate Education: Are We Even Close? Change 33 no 3 18-23 My/Je 2001

Tang, Thomas Li-Ping, Tang, David Shin-Hsiung and Tang, Cindy Shin-Yi (2004), College Tuition and Perceptions of Private University Quality, International Journal of Education Management, Emerald Group Publishing, Vol. 18, No. 5 – 2004, pp 304-316

Wachob, William (2009), An Investigation of Technology Funding in U.S. Colleges and Universities, (unpublished doctoral dissertation, Lawrence Technological University)

Appendix – I   **Raw Data**

rScore25	Admissions Rate	Endowment	Tuition and Fees (In St/Res)	Average 9 mo Faculty Salary
672.7	14.3	2,147,483,647.0	44,600	108,104
600.3	18.8	1,300,081,000.0	42,624	90,073
504.9	50.7	293,982,704.0	30,334	64,282
483.6	45.7	1,411,813,000.0	42,478	88,821
459.5	67.5	99,641,797.0	26,300	65,154
396.0	74.5	56,687,176.0	40,632	66,827
391.2	77.4	203,825,000.0	43,226	73,735
365.7	63.1	123,593,848.0	29,294	61,662
363.4	67.6	155,589,000.0	34,675	70,745
352.0	80.1	293,281,661.0	32,586	67,907
332.2	63.5	111,313,542.0	27,590	56,091
331.8	56.9	17,097,792.0	16,578	58,268
321.3	78.8	555,365,026.0	15,730	73,685
310.0	67.7	47,520,162.0	11,495	60,808
308.0	96.2	265,649,102.0	16,549	54,111
296.0	61.5	78,358,300.0	18,412	51,730
294.0	54.2	2,321,820.0	13,228	50,780
287.7	80.9	4,929,041.0	19,435	36,756
282.0	61.6	6,536,475.0	15,354	57,342
275.5	76.5	48,696,961.0	35,360	66,646
269.8	66.7	44,509,574.0	26,978	44,544
262.2	90.1	64,280,276.0	19,142	59,241
261.2	74.7	1,665,232.0	14,968	55,967
260.3	52.2	21,735,726.0	26,161	66,008
260.0	84.6	17,035,047.0	26,280	45,982
246.6	75.6		21,152	55,201
246.6	76.3	9,343,165.0	14,403	55,586
237.6	59.5	21,079,909.0	14,810	31,471
237.6	74.6	2,493,000.0	15,945	47,082
231.2	89.7	50,579,041.0	13,501	53,810
228.0	63.1	23,636,920.0	24,060	33,549
222.7	59.4	15,324,126.0	29,461	38,137
217.6	99.6	83,065,650.0	12,560	52,508
210.8	66.1	6,919,757.0	30,540	39,509
198.4	67.6	2,378,919.0	16,792	69,824
192.0	98.4		14,022	26,287
160.2	75.2	9,270,773.0	20,880	41,868

Table 6 - Raw Data

Appendix II **Data Graphs**
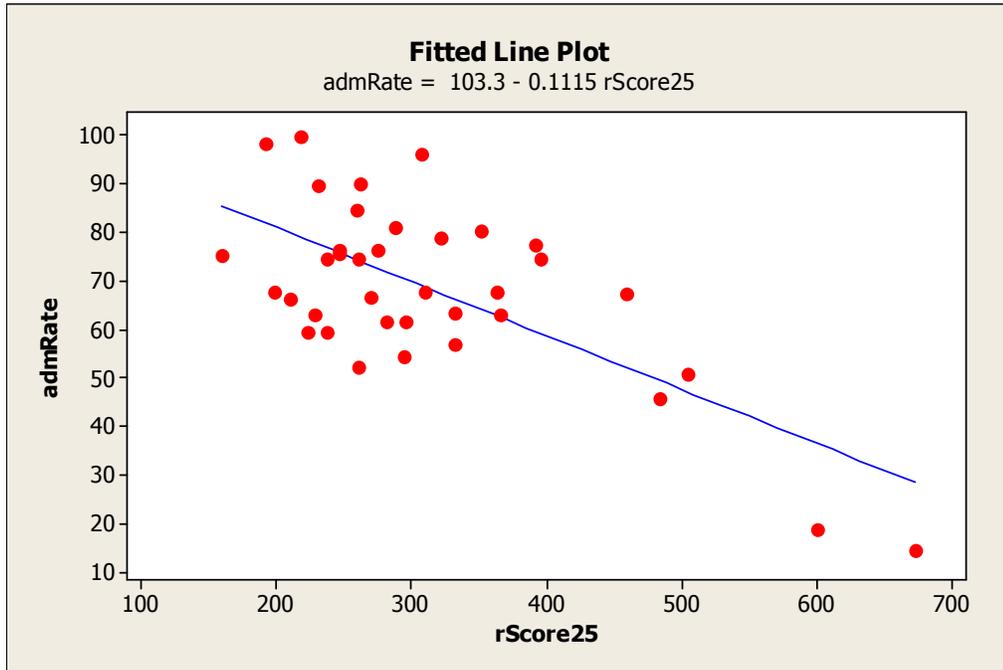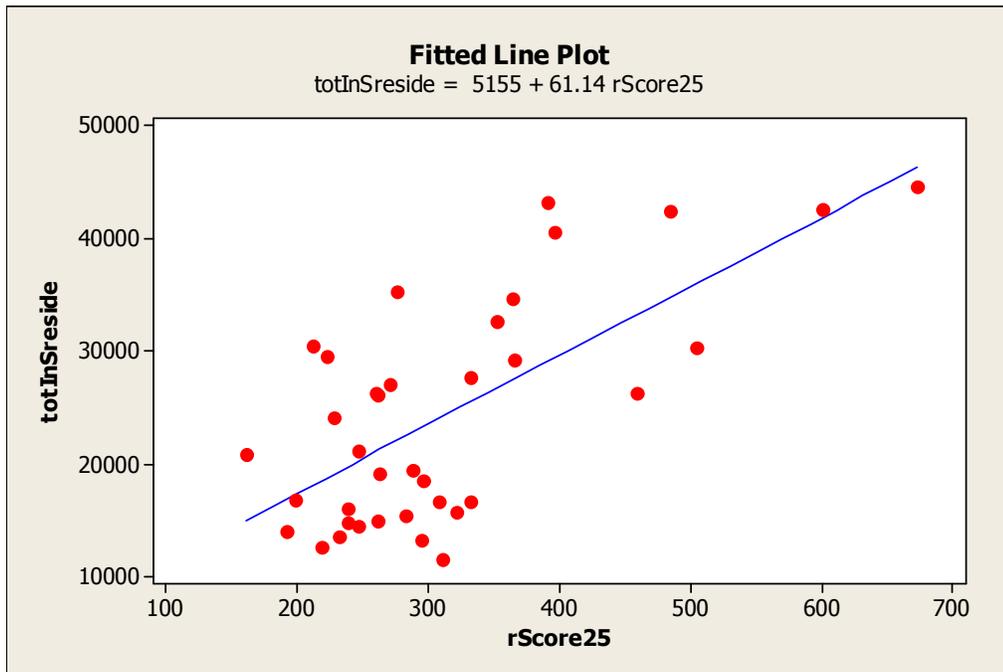

Figure 2 - Admissions Rate

Figure 3 - Faculty Salaries

Fitted Line Plot

totInSreside = 5155 + 61.14 rScore25
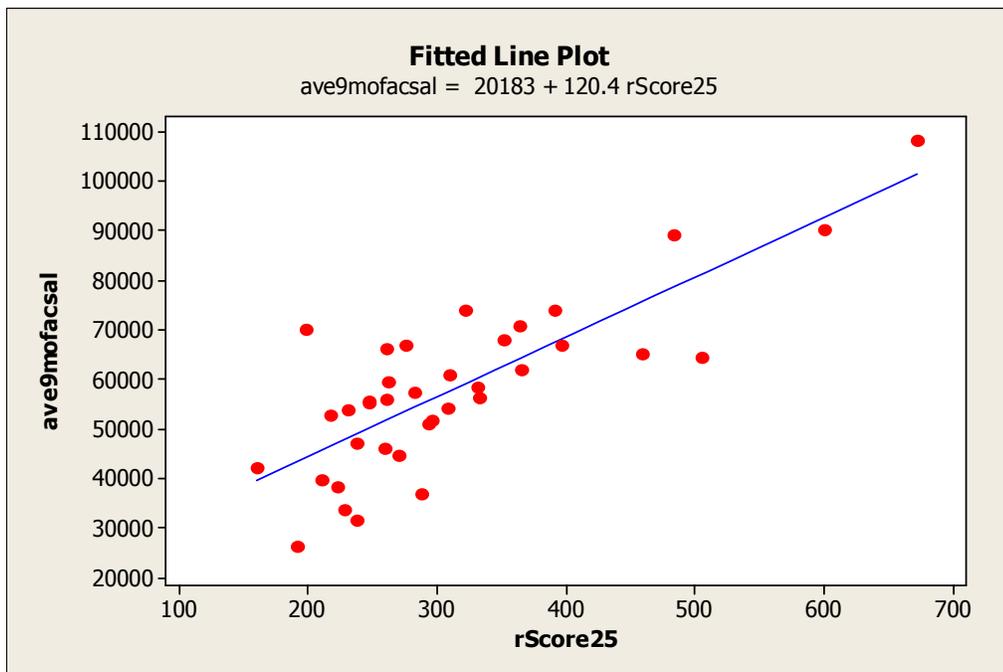
Figure 4 - Reported Tuition

Tuition reported for an in state, residential student

Fitted Line Plot

Log10-Endow = 5.914 + 0.005458 rScore25

S	0.546839
R-Sq	56.4%
R-Sq(adj)	55.1%
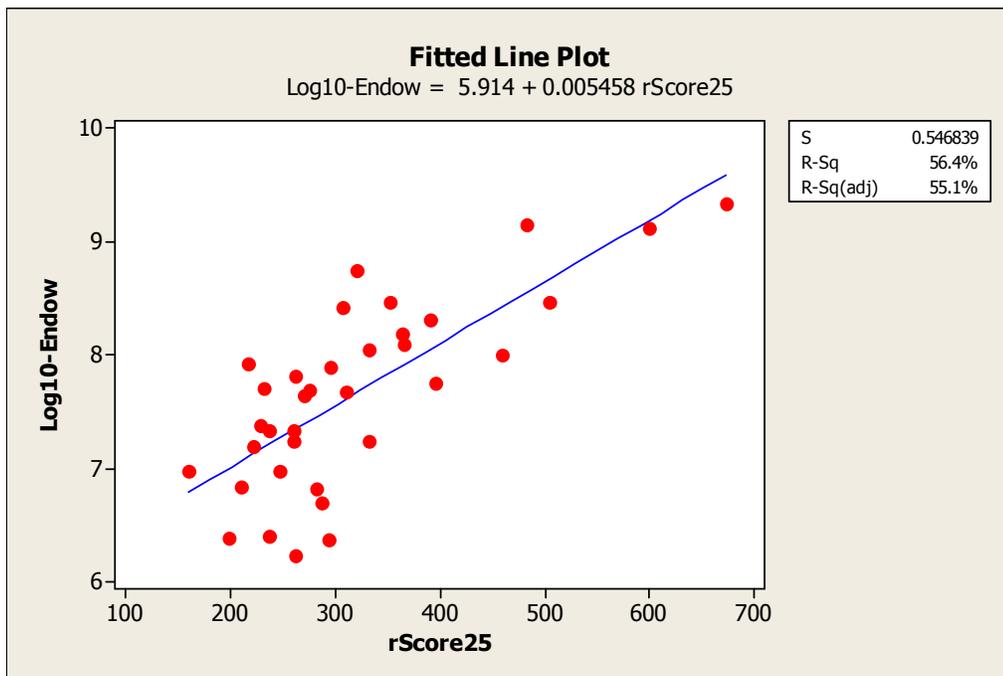
Figure 5 - Endowed Assets (log scaled)

Fitted Line Plot

endow = 6.48E+08 - 5184893 rScore25
+ 10724 rScore25**2

S	201621335
R-Sq	81.4%
R-Sq(adj)	80.3%
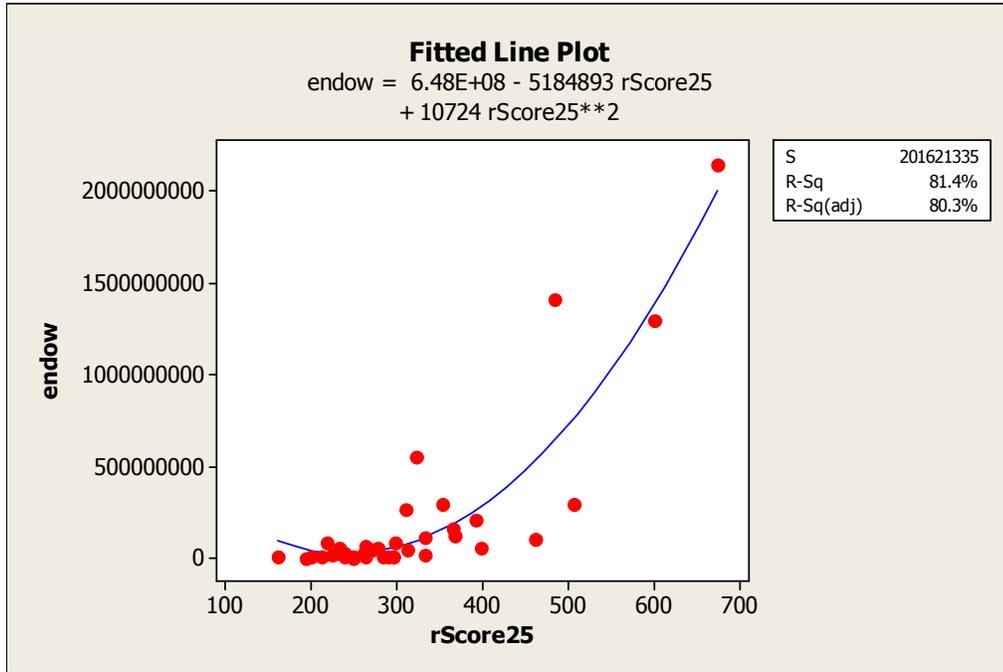
Figure 6 - Endowed Assets (quadratic)
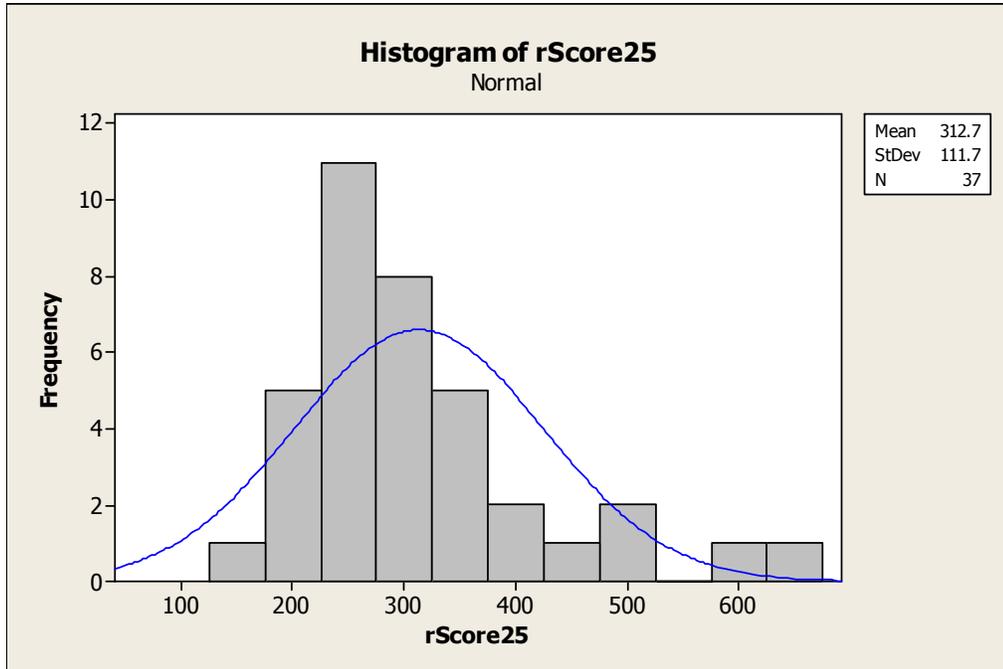
Histogram of rScore25

Normal

Mean	312.7
StDev	111.7
N	37

Figure 7 - Distribution of rScore values