James Lovell

A Statistical Analysis of SAT Scores

For my project, my criterion variable will be total SAT scores and my predictor variable will be the wealth of their school district (District Wealth per Student).

SAT scores are of large economic significance, as they determine much of a young person's future. The difference of just a few points could be the difference in a young person being able to go to higher education or not.

It is likely that there will be a positive relationship between SAT scores and wealth of a district, as wealthier districts are more able to pay for SAT prep courses.

I have chosen to sort my data into subgroups based on the enrolment

I plan on sorting school's enrolment into ascending order, then splitting that list into 4 different groups. It is possible that schools with higher or lower enrolment perform differently on the SAT, so it would be economically significant to split the data set into groups to see if there is any inequality in that.

Schools with smaller enrolment can focus more on individual students, and therefore will likely have higher SAT scores.

I believe smaller schools and bigger schools will have different scores, but I believe that the degree to which their SAT scores are improved by additional wealth is the same as one another

I believe that the subgroups with lower enrolment will have higher SAT scores.

Literature Review

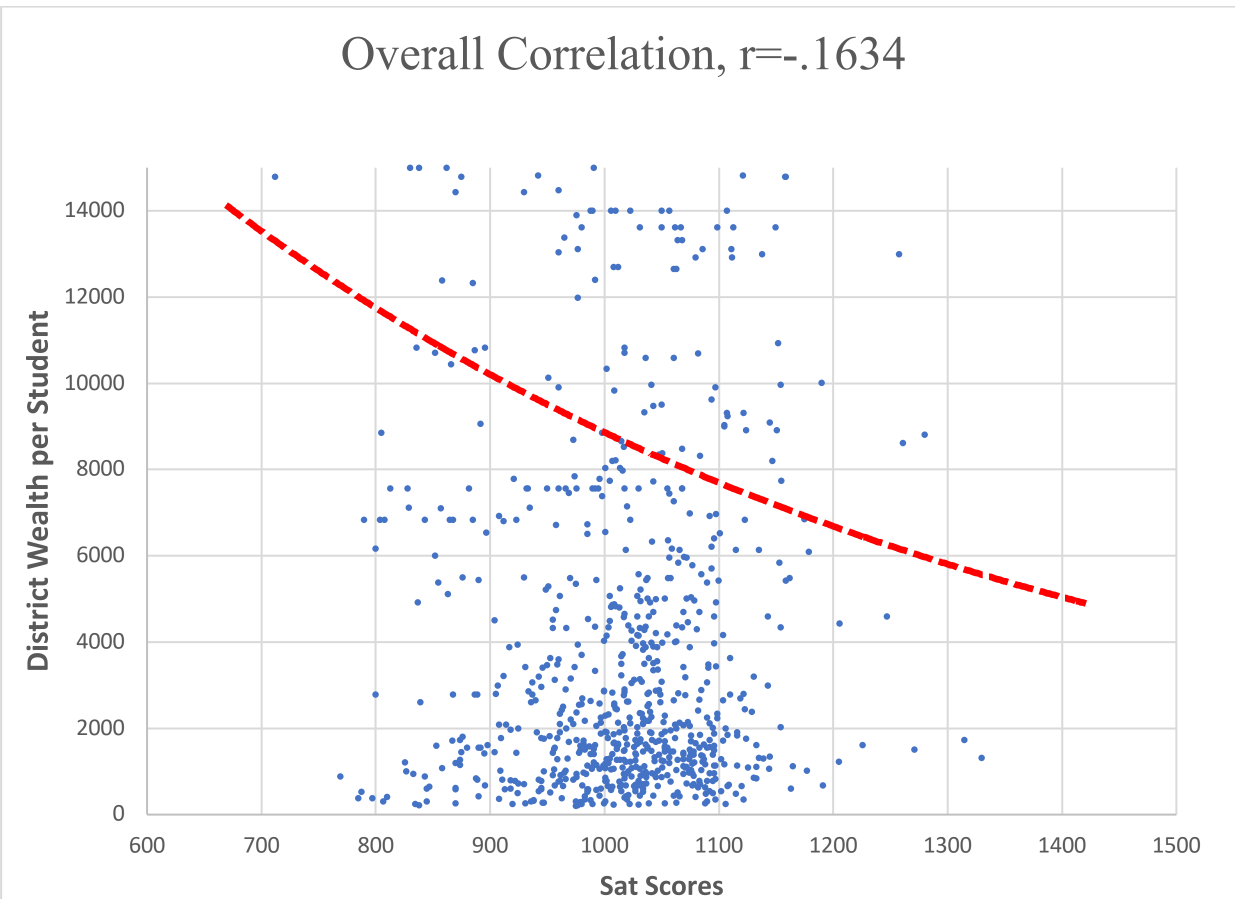
In reference to my criterion and predictor variables, In a magazine article in "The Phi Delta Kappan", a professional magazine covering topics of anything k-12, the author states "The NCES researchers computed the percentages of students who met each of several academic standards intended to resemble the admissions criteria of selective colleges. They compared the results for students of high, medium, and low socioeconomic status (SES), as measured by parental education, occupation, and income. Thirty-two percent of the high-SES students earned a combined SAT score of at least 1100, but only 9% of the low-SES group did so."[1] (Zwick, 2002). This research that would strongly suggest that wealth is tied to SAT scores. This is just one of many studies that points to this conclusion.

Descriptive Statistics

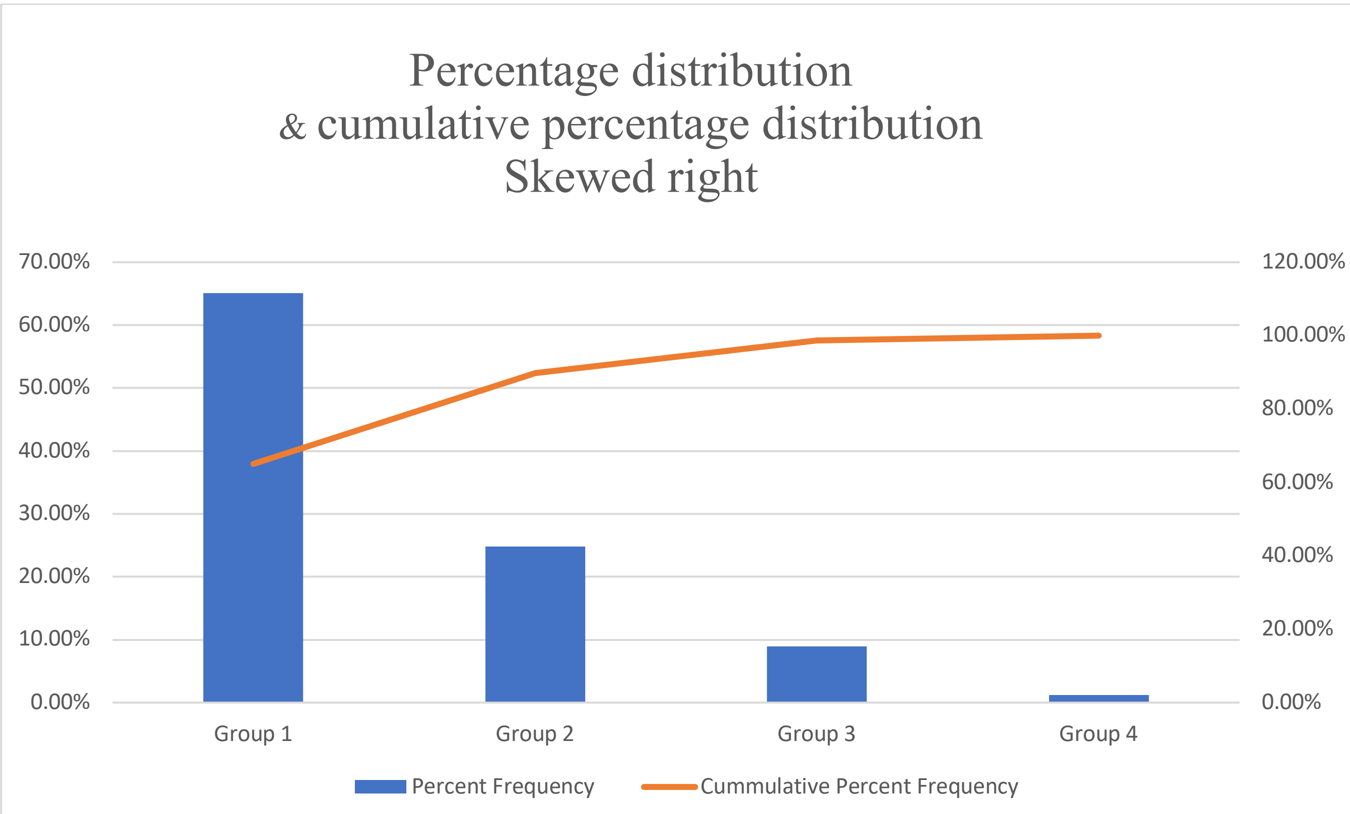
Scores	Overall	Group 1	Group 2	Group 3	Group 4
Minimum	669	669	810	831	885
Maximum	1421	1421	1280	1280	1207
Range	752	752	470	449	322
Count	1282	834	318	114	16
Mean	1011.294	1008.643	1008.440	1036.860	1024.063
Median	1019	1020	1011.500	1024.500	1031.500

Scores	Overall	Group 1	Group 2	Group 3	Group 4
Sample Var.	9880.247	9642.789	9815.061	11252.900	10375.263
Standard Dev.	99.399	98.198	99.071	106.080	101.859

Correlation	Overall	Group 1	Group 2	Group 3	Group 4
r	-0.163	-0.143	-0.371	-0.085	0.021



Frequency Distribution	SAT Scores	Count	Percent Freq.	Cumulative Percent Freq.
Group 1	34-1302.75	834	65.05%	65.05%
Group 2	1302.75-2571.5	318	24.80%	89.86%
Group 3	2571.5-3840.25	114	8.89%	98.75%
Group 4	3840.25-5109	16	1.25%	100.00%



- The biggest discrepancy that is observed is the heavy right skew of the group's distributions. This shows that a very large amount of schools has a smaller enrolment, with a couple outliers with very large enrolment.

Single Sample Confidence Intervals and Hypothesis Tests

Mean Confidence	Overall	Group 1	Group 2	Group 3	Group 4
UCL 99%	1018.46	1017.42	1022.84	1062.89	1099.10
UCL 95%	1016.74	1015.32	1019.37	1056.54	1078.34
UCL 90%	1015.86	1014.24	1017.61	1053.34	1068.70
Mean	1011.29	1008.64	1008.44	1036.86	1024.06
LCL 90%	1006.72	1003.04	999.28	1020.38	979.42
LCL 95%	1005.85	1001.97	997.51	1017.18	969.79
LCL 99%	1004.13	999.86	994.04	1010.83	949.03
Var Confidence	Overall	Group 1	Group 2	Group 3	Group 4
UCL 99%	10954.98	10964.92	12116.03	16124.14	30265.02
UCL 95%	10683.71	10626.84	11502.64	14727.14	22529.89
UCL 90%	10548.36	10459.21	11204.64	14073.74	19547.33
Mean	9880.25	9642.79	9815.06	11252.90	10375.26
LCL 90%	9262.66	8902.23	8628.63	9087.84	5918.30
LCL 95%	9150.72	8769.77	8425.18	8742.89	5395.29
LCL 99%	8937.12	8518.39	8045.76	8117.98	4541.63

Sub-group Mean vs. Overall Mean

T Statistics for Mean	Group 1	Group 2	Group 3	Group 4
t-calc	-0.78	-0.51	2.57	0.50
t-crit @ 90%	1.28	1.28	1.29	1.34
t-crit @ 95%	1.65	1.65	1.66	1.75
t-crit @ 99%	2.33	2.34	2.36	2.60
Hypothesis Results	Group 1	Group 2	Group 3	Group 4
90% Confidence	Fail to Reject	Fail to Reject	Reject	Fail to Reject
95% Confidence	Fail to Reject	Fail to Reject	Reject	Fail to Reject
99% Confidence	Fail to Reject	Fail to Reject	Reject	Fail to Reject

- This is to determine if the means of our individual sub-groups are different from that of our "overall" mean. In group 1, 2, and 4 we can see that we are unable to definitively say they are different from our overall mean, at any of our three confidence levels, so they are said to be "Failed to Reject".

Sub-group Variance vs. Overall Variance

X Statistic at 90%	Group 1	Group 2	Group 3	Group 4
X Crit Upper	885.718	349.667	132.643	22.307
X Calc	812.980	314.909	128.699	15.752
X Crit Lower	781.138	285.190	94.213	8.547
Conclusion	Fail to Reject	Fail to Reject	Fail to Reject	Fail to Reject

X Statistic at 95%	Group 1	Group 2	Group 3	Group 4
X Crit Upper	901.26	359.52	138.81	25.00
X Calc	812.98	314.91	128.70	15.75
X Crit Lower	767.02	276.75	89.46	7.26
Conclusion	Fail to Reject	Fail to Reject	Fail to Reject	Fail to Reject

X Statistic at 99%	Group 1	Group 2	Group 3	Group 4
X Crit Upper	930.88	378.50	150.88	30.58
X Calc	812.98	314.91	128.70	15.75
X Crit Lower	741.00	261.38	80.99	5.23
Conclusion	Fail to Reject	Fail to Reject	Fail to Reject	Fail to Reject

- This is a test of if the sample variances of our individual groups are the same as our overall sample (population) variance. All of my hypothesis test for every confidence level are “Fail to Reject” which means that we are unable to say that any of my sub-groups variances are <u>different</u> than the overall.

Two Sample Confidence Intervals and Hypothesis Tests

Mean 95%	Group 1 & 2	Groups 1 & 3	Groups 1 & 4	Group 2 & 3	Group 2 & 4	Group 3 & 4
UCL	12.932	-8.783	33.258	-6.757	34.374	68.576
LCL	-12.527	-47.651	-64.097	-50.081	-65.619	-42.982
Tcrit	1.962	1.977	2.120	1.973	2.120	2.086
T calc	0.031	-2.687	-0.600	-2.497	-0.599	0.468
Conclusion	Fail to Reject	Reject	Fail to Reject	Reject	Fail to Reject	Fail to Reject

- My pairwise groups conclusions are not all the same, but I have 4 “Fail to Reject” and 2 “Reject”, so the majority of my pairwise groups means are close together.

Pooled Variances to Calculate Confidence Intervals	Calculated
Pooled variance 1&2	9690.276
Pooled variance 1&3	9835.118
Pooled variance 1&4	9655.746
Pooled variance 2&3	10192.912
Pooled variance 2&4	9840.371
Pooled variance 3&4	11150.052

- This data was used to find the upper and lower confidence intervals for the difference of the means above.

Variance 95%	Group 1 & 2	Groups 1 & 3	Groups 1 & 4	Group 2 & 3	Group 2 & 4	Group 3 & 4
F calc	1.018	1.167	1.076	1.146	1.057	1.085
F crit	1.170	1.279	2.073	1.303	2.084	2.117
Conclusion	Fail to Reject	Fail to Reject	Fail to Reject	Fail to Reject	Fail to Reject	Fail to reject

- The results are “Fail to Reject” for all the tests of difference of variance between the different sub-groups, which means that the variances is relatively uniform across my pairwise groups.

ANOVA Tests

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Group 1	834	841208	1008.643	9642.789
Group 2	318	320684	1008.440	9815.061
Group 3	114	118202	1036.860	11252.900
Group 4	16	16385	1024.063	10375.263

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	85571.558	3	28523.853	2.900	0.034	2.612
Within Groups	12571024.6	1278	9836.482			
Total	12656596.1	1281				

- Our F calc > F crit, so we would reject this. This means that our means are not equal across our different sub-groups.

Correlation

Correlation between groups	1 & 2	1 & 3	1 & 4	2 & 3	2 & 4	3 & 4
r	0.045	-0.018	0.031	0.109	-0.134	0.022

Correlation Analysis	Sample Corr. Coeff.	Sample Sizes	zr	n-3		
Group 1	-0.143	834	-0.144	831	17.263	-119.774
Group 2	-0.371	318	-0.390	315	47.799	-122.705
Group 3	-0.085	114	-0.085	111	0.805	-9.453
Group 4	0.021	16	0.021	13	0.006	0.272
				1270	65.873	-251.661

Hypothesis test at 95%	Calculated
Chi-square calc	16.004
Critical (upper) Chi-Square for α=5%	9.488
Conclusion	Reject

- Our rejection of this hypothesis means that there is no significant correlation between our predictor variable and our criterion variable.

Conclusion:

Through this testing, we have found no significant correlation between our predictor variable and our criterion variable. In real world terms this means that, according to our tests, using the data provided, there is not significant proof that the wealthier the school district is, the higher a students SAT score will be. In addition to this, my chosen sub-groups (school enrolment) also proved to have little effect on the criterion variable. This was shown through the majority of our hypothesis tests comparing sub-groups to one another resulting in “Fail to Reject” which means that there was no significant difference between our different subgroups.

One problems that could cause issue with my sub-groups is that their sample sizes are very different, with my largest sub-group engulfing ~65% of my observations, and the smallest sub-group only capturing ~1% of my observations. Another issue is that “District Wealth Per Student” might not be an accurate way to measure how much money the families of kids attending school have. District wealth per student is defined as “A measure of the School District’s taxable property per student” which is quite different than a measure of the wealth of school children’s families.

References

1. Rebecca Zwick. (2002). Is the SAT a “Wealth Test”? The Phi Delta Kappan, 84(4), 307. Retrieved from http://search.ebscohost.com.srv-proxy2.library.tamu.edu/login.aspx?direct=true&db=edsjsr&AN=edsjsr.20440340&site=eds-live