Eric Douglas

UCLA UP220B, W09

QUANTITATIVE ANALYSIS OF LOS ANGELES HIGH SCHOOLS

"Educational Accountability"

In this exercise, I sought to find linkages between academic performance in Los Angeles high schools and demographic characteristics of students attending them. To do this, I used the 2005 Academic Performance Index (API) database (http://api.cde.ca.gov) records for the roughly 1,000 LA high schools on file. The API scores range from 200-1000 and the database includes information on various demographic categories.

My initial intuition was to compare API scores ("API05") against the categories "MEALS" (percentage of students tested that are participants in the free or reduced price lunch program), "EL" (percentage English language learners), "ACS_CORE" (number of core academic courses), "AVG_ED" (average parent education level), and "FULL" (percent teachers at this school with full credentials). I found a high level of correspondence among API05, MEALS, EL, and AVG_ED. I interpreted this to mean that lower scores had a connection to poorer areas and areas with students learning English, and higher scores had a connection to households with college-educated parents.

Then I looked at Appendix D in the "2005 Supplement to PSAA Technical Report 00-1" which listed bivariate correlations among the 1034 high schools in LA County with respect to API scores. From this, I noticed that, in addition to parental educational level, English-speaking ability, and assumed poverty level, API scores held significant

correspondence with both Hispanic (negative) and White (positive) demographics. I found it interesting that the African American demographic held only a moderate (negative) correspondence, since it tends to suffer from similar societal challenges to those of the Hispanic population. Why was there such a strong negative connection between the Hispanic demographic and school performance? It may be that, in Los Angeles, they have a combination of challenges such as recent immigration, language difficulty, poverty, and legal status. Even though other ethnic groups may share these difficulties, they may simply not have the numbers to affect school performance.

As I was reviewing a scatterplot of parental educational level against API score, I noticed a seeming miracle. The connection between the two was so strong, but it seemed that a few schools had beaten the system. They had tremendous API scores, even though their parental educational level was as low as could be. What had these schools done to score so well despite the odds? What programs did they have in place to overcome this formidable educational obstacle? Why were they succeeding where all others had failed?

I took a closer look. It turned out that all of these miracle schools were within two school districts—Irvine Unified and Capistrano Unified in Orange County. Obviously, some visionary district superintendent had come up with an amazing educational program to overcome the disadvantage of low parental educational level. So, I looked at the district-level numbers. Then, I looked at the school-level numbers, and discovered the miracle. They lied. It turns out, the lowest level of education counts as zero. If you don't report, it also counts as zero. Neither Irvine nor Capistrano reported parental educational level. So, they looked like heroes. Unfortunately, they were just

outliers (out-liars?). When I removed them from the equation, the correspondence between parental educational level and API score grew even closer.

Getting back on track, I wanted to figure out what factors contributed to high or low API scores. The two big positive factors were Whiteness and education. The negative categories were Hispanic, English-learners, and reduced meal people. Unfortunately, the negative categories had a high level of colinearity, so it was hard to tell which of the three was most influential. My response was to take a second look at the positive factor of education and the negative factor of poverty. As I looked at the individual parental education categories, I noticed that the lower categories, no high school and high school grad, were about the same. Also, the upper categories, college grad and graduate school grad were about the same. So, I focused on the correspondence between college grad and API. My conclusion, then, is that there is a high correspondence among parental college graduation, poverty, and API score. So, the take-home point is, graduate from college and avoid poverty. Our schools are counting on you.

Correlations

		api05	meals	acs_core	el	avg_ed	full
api05	Pearson Correlation	1	-.619**	.087**	-.557**	.648**	.326**
	Sig. (2-tailed)		.000	.006	.000	.000	.000
	N	1047	1047	996	1047	1047	1047
meals	Pearson Correlation	-.619**	1	-.042	.687**	-.688**	-.276**
	Sig. (2-tailed)	.000		.182	.000	.000	.000
	N	1047	1047	996	1047	1047	1047
acs_core	Pearson Correlation	.087**	-.042	1	.022	.001	.053
	Sig. (2-tailed)	.006	.182		.494	.972	.094
	N	996	996	996	996	996	996
el	Pearson Correlation	-.557**	.687**	.022	1	-.630**	-.239**
	Sig. (2-tailed)	.000	.000	.494		.000	.000
	N	1047	1047	996	1047	1047	1047
avg_ed	Pearson Correlation	.648**	-.688**	.001	-.630**	1	.222**
	Sig. (2-tailed)	.000	.000	.972	.000		.000
	N	1047	1047	996	1047	1047	1047
full	Pearson Correlation	.326**	-.276**	.053	-.239**	.222**	1
	Sig. (2-tailed)	.000	.000	.094	.000	.000	
	N	1047	1047	996	1047	1047	1047

**. Correlation is significant at the 0.01 level (2-tailed).

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.748[a]	.559	.557	57.619

a. Predictors: (Constant), full, acs_core, avg_ed, el, meals

b. Dependent Variable: api05

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4174031.137	5	834806.227	251.456	.000[a]
	Residual	3286693.211	990	3319.892		
	Total	7460724.348	995			

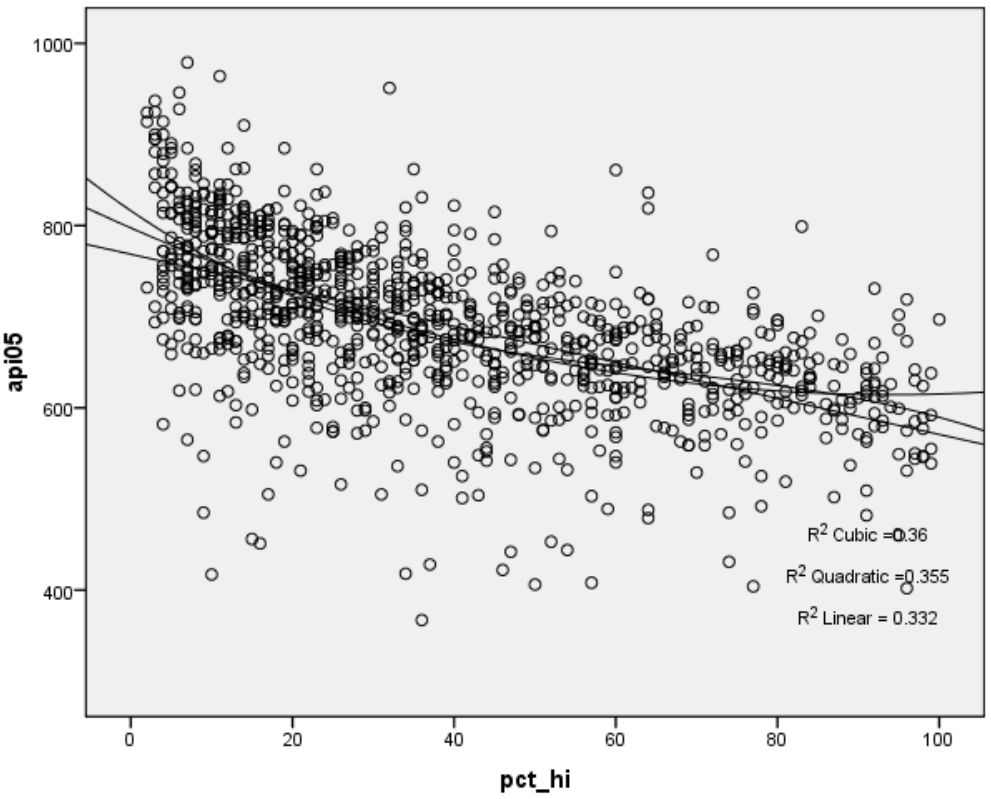
a. Predictors: (Constant), full, acs_core, avg_ed, el, meals

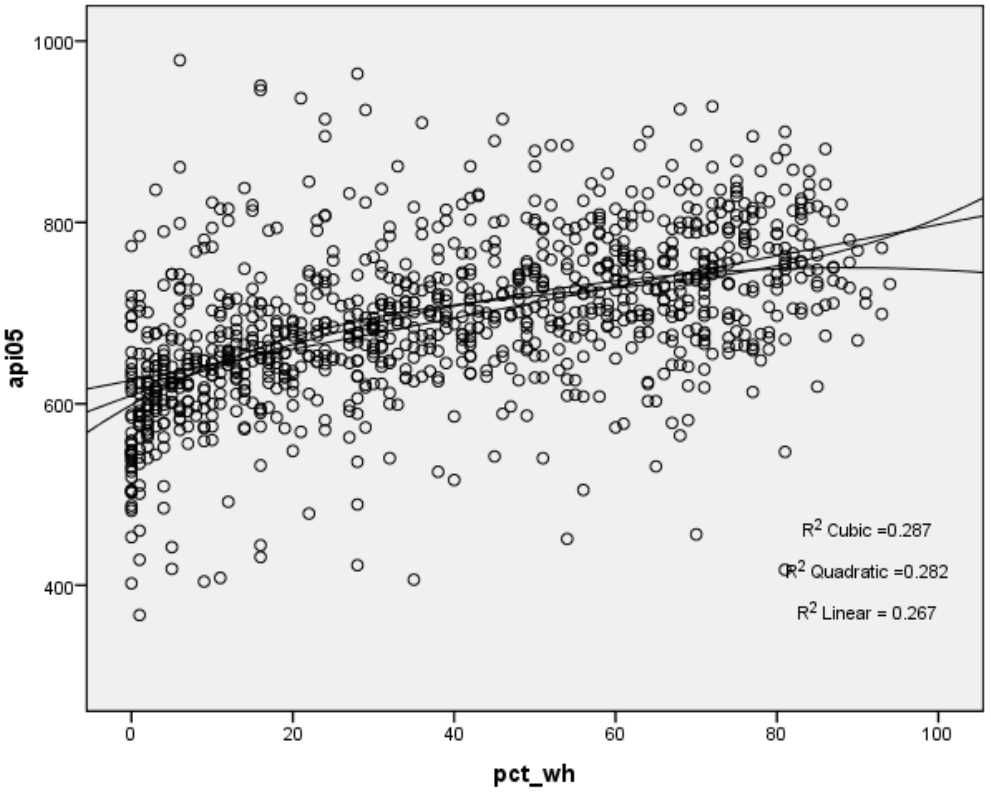
b. Dependent Variable: api05

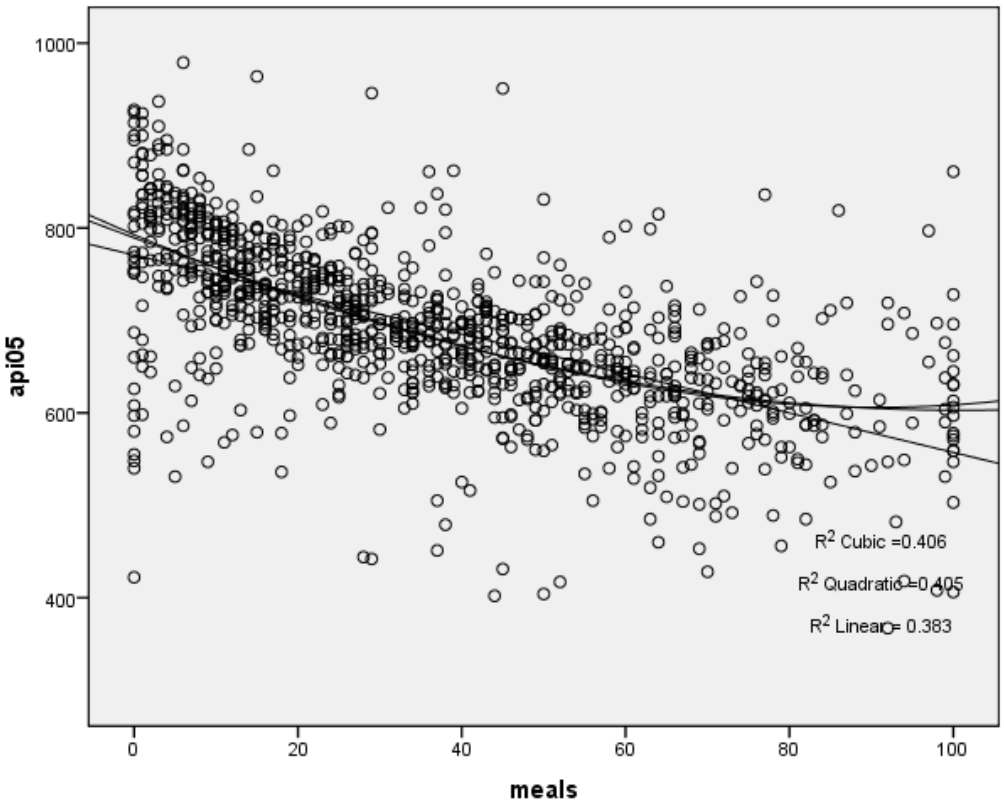
Coefficients[a]

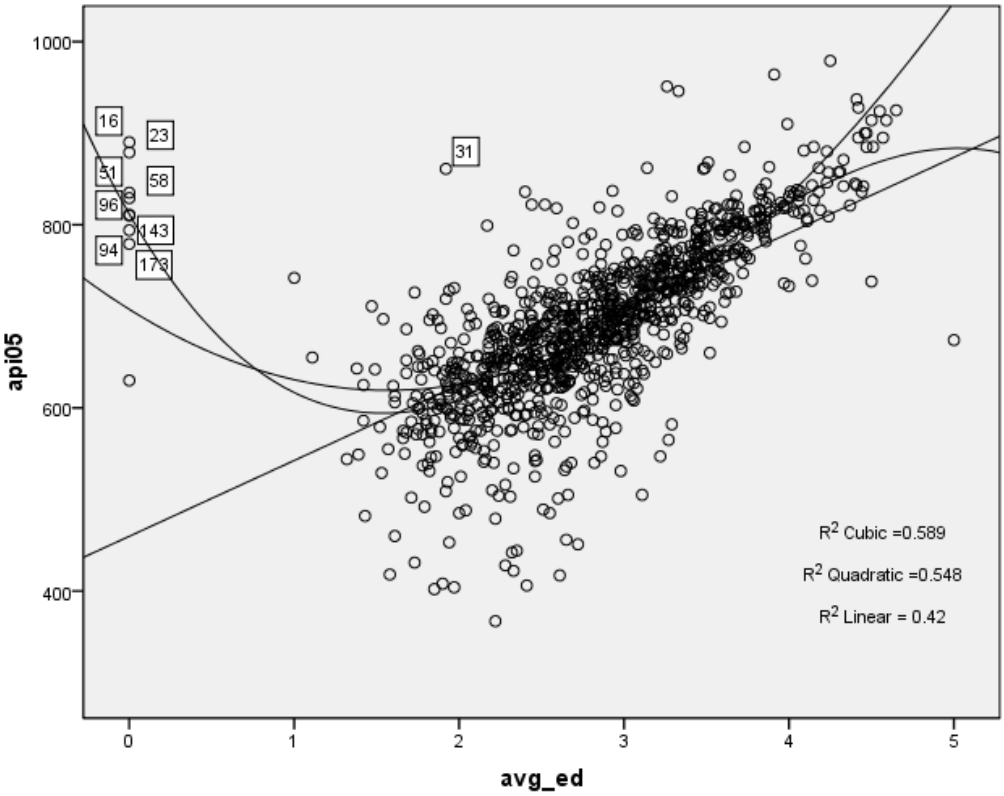
Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	517.494	23.903		21.649	.000
	meals	-.825	.113	-.245	-7.280	.000
	el	-1.079	.178	-.191	-6.066	.000
	acs_core	1.514	.426	.075	3.550	.000
	avg_ed	44.245	3.846	.360	11.505	.000
	full	.689	.166	.092	4.146	.000

a. Dependent Variable: api05

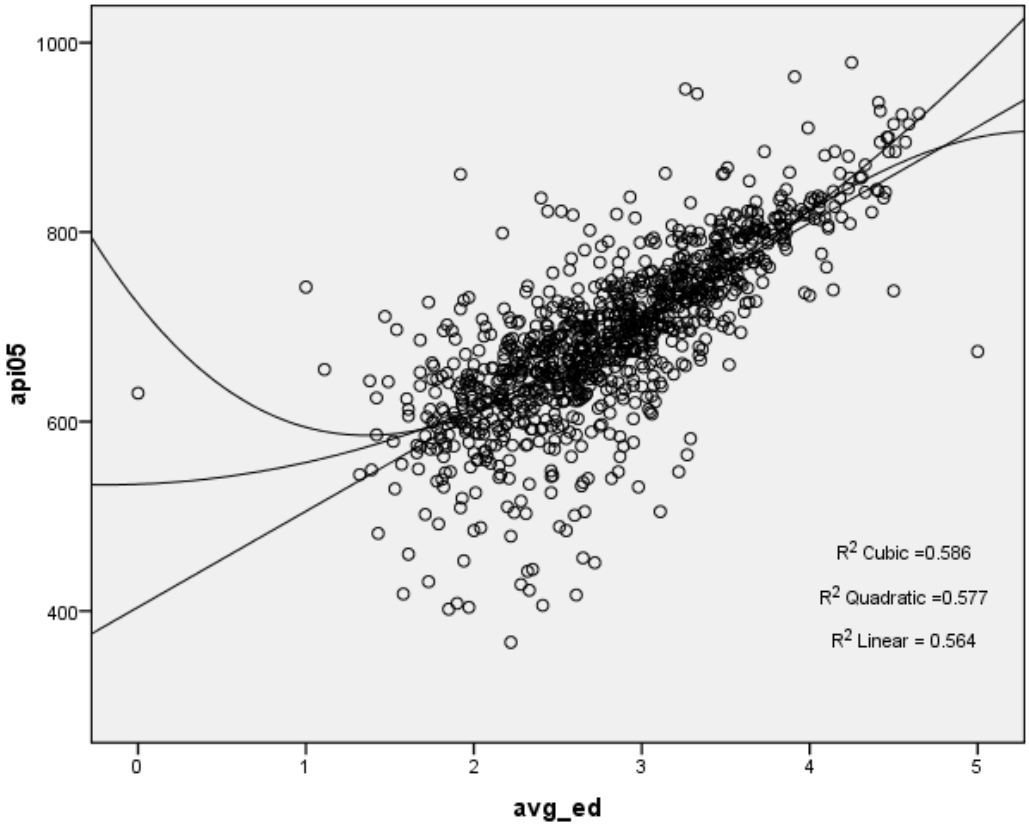

1000
800
600
400
api05
0
20
40
60
80
100
pct_hi
R2 Cubic =0.36
R2 Quadratic =0.355
R2 Linear = 0.332


1000
800
600
400
api05
0
20
40
60
80
100
pct_wh
R2 Cubic =0.287
R2 Quadratic =0.282
R2 Linear = 0.267


1000
800
600
400
api05
0
20
40
60
80
100
meals
R2 Cubic =0.406
R2 Quadratic =0.405
R2 Linear = 0.383


1000
800
600
400
api05
16
23
31
51
58
96
143
94
173
0
1
2
3
4
5
avg_ed
R2 Cubic =0.589
R2 Quadratic =0.548
R2 Linear = 0.42

FILTER OFF. USE ALL. SELECT IF (dname ~= "Irvine Unified" & dname ~= "Capistrano Unified"). EXECUTE. SORT CASES BY api05 (D). GRAPH /SCATTERPLOT(BIVAR)=avg_ed WITH api05 /MISSING=LISTWISE.



Correlations

		api05	pct_wh	pct_hi	el	meals	avg_ed
api05	Pearson Correlation	1	.514**	-.571**	-.557**	-.614**	.751**
	Sig. (2-tailed)		.000	.000	.000	.000	.000
	N	1038	1038	1038	1038	1038	1038
pct_wh	Pearson Correlation	.514**	1	-.764**	-.683**	-.673**	.665**
	Sig. (2-tailed)	.000		.000	.000	.000	.000
	N	1038	1038	1038	1038	1038	1038
pct_hi	Pearson Correlation	-.571**	-.764**	1	.758**	.712**	-.783**
	Sig. (2-tailed)	.000	.000		.000	.000	.000
	N	1038	1038	1038	1038	1038	1038
el	Pearson Correlation	-.557**	-.683**	.758**	1	.686**	-.692**
	Sig. (2-tailed)	.000	.000	.000		.000	.000
	N	1038	1038	1038	1038	1038	1038
meals	Pearson Correlation	-.614**	-.673**	.712**	.686**	1	-.777**
	Sig. (2-tailed)	.000	.000	.000	.000		.000
	N	1038	1038	1038	1038	1038	1038
avg_ed	Pearson Correlation	.751**	.665**	-.783**	-.692**	-.777**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	
	N	1038	1038	1038	1038	1038	1038

**. Correlation is significant at the 0.01 level (2-tailed).

[DataSet1] H:\220b\final\douglas-schoolapi05V2.sav

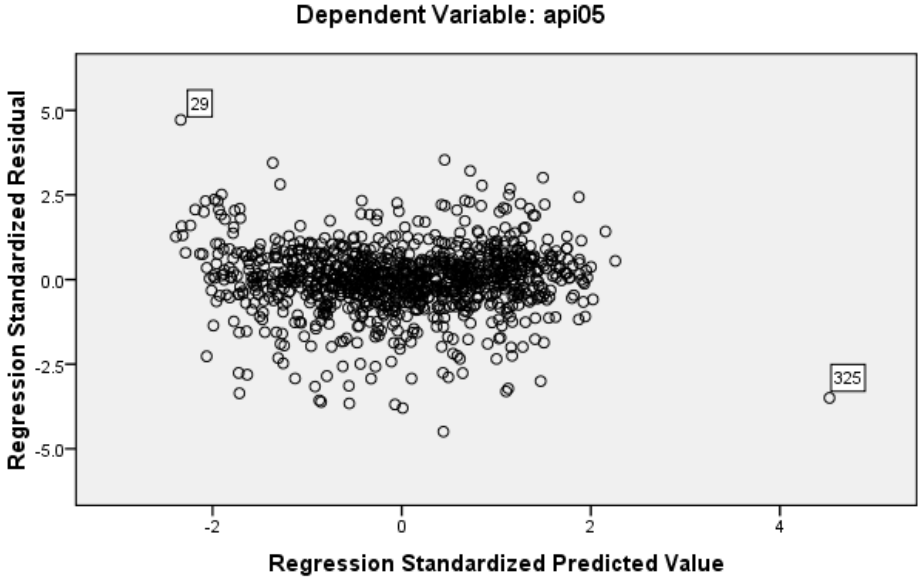
Correlations

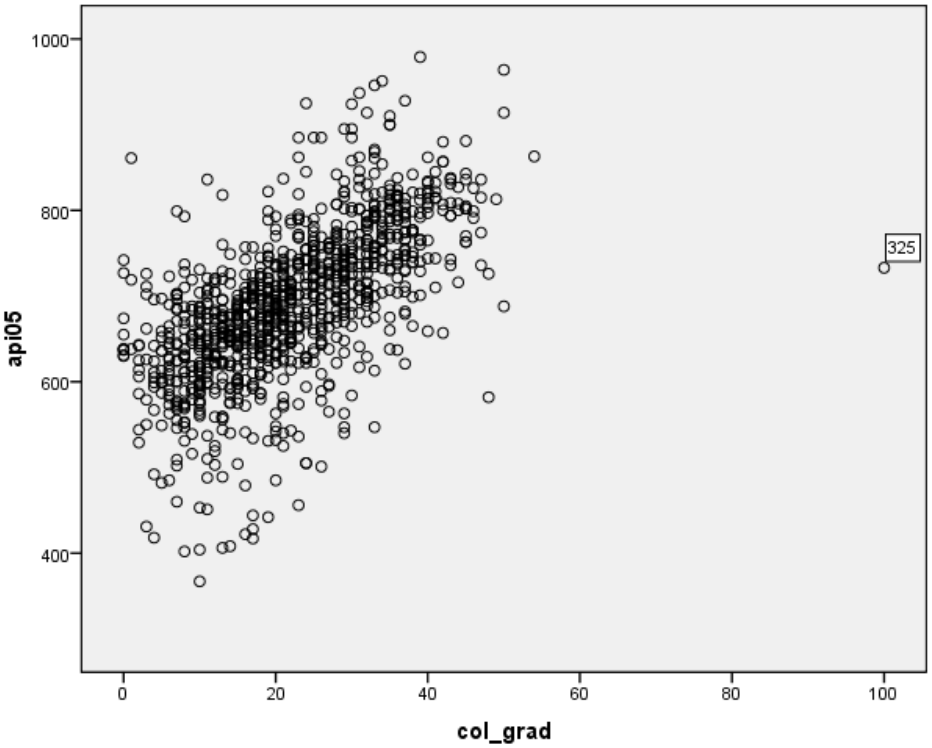
		api05	not_hsg	hsg	some_col	col_grad	grad_sch
api05	Pearson Correlation	1	-.607**	-.597**	.102**	.624**	.689**
	Sig. (2-tailed)		.000	.000	.001	.000	.000
	N	1038	1038	1038	1038	1038	1038
not_hsg	Pearson Correlation	-.607**	1	.332**	-.464**	-.747**	-.563**
	Sig. (2-tailed)	.000		.000	.000	.000	.000
	N	1038	1038	1038	1038	1038	1038
hsg	Pearson Correlation	-.597**	.332**	1	-.084**	-.588**	-.696**
	Sig. (2-tailed)	.000	.000		.007	.000	.000
	N	1038	1038	1038	1038	1038	1038
some_col	Pearson Correlation	.102**	-.464**	-.084**	1	.076*	-.160**
	Sig. (2-tailed)	.001	.000	.007		.014	.000
	N	1038	1038	1038	1038	1038	1038
col_grad	Pearson Correlation	.624**	-.747**	-.588**	.076*	1	.551**
	Sig. (2-tailed)	.000	.000	.000	.014		.000
	N	1038	1038	1038	1038	1038	1038
grad_sch	Pearson Correlation	.689**	-.563**	-.696**	-.160**	.551**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	
	N	1038	1038	1038	1038	1038	1038

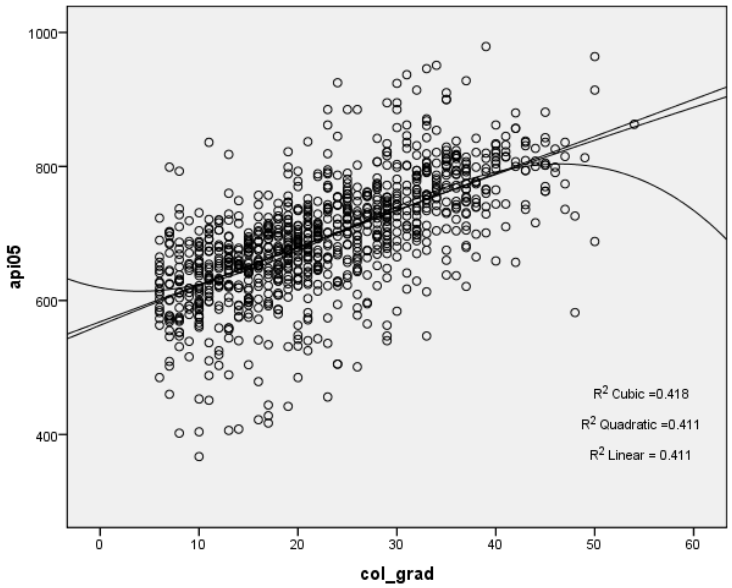
**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Scatterplot







USE ALL. COMPUTE filter_$=(col_grad > 5 & col_grad 95 & meals > 5 & meals 95). VARIABLE LABEL filter_$ 'col_grad > 5 & col_grad 95 & meals > 5 & meals 95 (FILTER)'. VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'. FORMAT filter_$ (f1.0). FILTER BY filter_$. EXECUTE. REGRESSION /MISSING LISTWISE /STATISTICS COEFF OUTS R ANOVA COLLIN TOL /CRITERIA=PIN(.05) POUT(.10) /NOORIGIN /DEPENDENT api05 /METHOD=ENTER col_grad meals /SCATTERPLOT=(*ZRESID ,*ZPRED) /RESIDUALS HIST(ZRESID).

Scatterplot

Dependent Variable: api05