SAT Scores Regression Interpretations

<u>The summary of my results for model_1 can be seen in the table below:</u>

	Coefficient Estimate	**Standard Errors**	**T-Stats**
Intercept	1752.44	30.72	57.04
rexppp	-0.01245	0.00232	-5.37
Model R^2	0.1251		
Model Adj. R^2	0.1207		

- The coefficient estimate can be interpreted as a $1000 increase in pupil real expenditures is predicted to reduce pupil SAT scores by 12.45 points.

<u>The summary of my results for model_2 can be seen in the table below:</u>

	Coefficient Estimate	**Standard Errors**	**T-Stats**
Intercept	1661.76	15.57	106.74
rexppp	0.00556	0.00135	4.13
pertake	-3.71998	0.14815	-25.11
Model R^2	0.7885		
Model Adj. R^2	0.7864		

- The coefficient estimate can now be interpreted as a $1,000 increase in per pupil real expenditures is predicted to raise SAT scores by 5.56 points, holding pertake constant. You can see from the decrease in this value that including PERTAKE in our model vastly reduced our final coefficient estimate. Our R^2 value also vastly improved, suggesting that including PERTAKE helps our model explain a much larger percentage of the variation in SAT scores.

<u>Summary of our Means Procedure to see if PERTAKE varies over time</u>

- Both Iowa and Utah showed no variation in PERTAKE over the 4 years of data.
- Only 12 states that had a coefficient of variation greater than 0.08, and I have listed them below:
    1.) Arkansas
    2.) Illinois
    3.) Kentucky
    4.) Michigan
    5.) Mississippi
    6.) Missouri
    7.) Nebraska

8.) North Dakota
9.) Oklahoma
10.) South Dakota
11.) Wisconsin
12.) Wyoming

<u>The summary of my results for model_3 can be seen in the table below:</u>

	Coefficient Estimate	Standard Errors	T-Stats
Intercept	1639.497	32.569	50.34
rexppp	0.00254	0.00183	1.39
Model R^2	0.9954		
Model Adj. R^2	0.9938		

- I ran an F-Test to see if the inclusion of the dummy variables helps improve the models using the formulas below:

$$F^* = \frac{\frac{(R_{UR}^2 - R_R^2)}{q}}{\frac{(1-R_{UR}^2)}{(n-k-1)}} \quad \rightarrow \quad F^* = \frac{\frac{(0.9954-0.1251)}{50}}{\frac{(1-0.9954)}{(204-50-1)}} \quad \rightarrow \quad F^* = 578.94$$

- You can see that such a high critical F value tells us that the inclusion of our dummy variables significantly improved our model
- The coefficient can now be interpreted as a $1,000 increase in per pupil expenditure is expected to increase student SAT scores by 2.54 points.

<u>The summary of my results for model_4 can be seen in the table below:</u>

	Coefficient Estimate	Standard Errors	T-Stats
Intercept	1673.16	29.457	56.80
rexppp	0.00131	0.00164	0.80
pertake	-2.2797	0.3563	-6.40
Model R^2	0.9964		
Model Adj. R^2	0.9951		

- Out of the four models, I would choose to report on model four. Although the rexppp variable is not significant at a 10% level – as seen through its p-value of 0.4260 on my output document, you can see that the R^2 value is the highest in this model. The increase in R^2 when we included the dummy variables suggest that the dummy variables removed the omitted variable bias in our model.