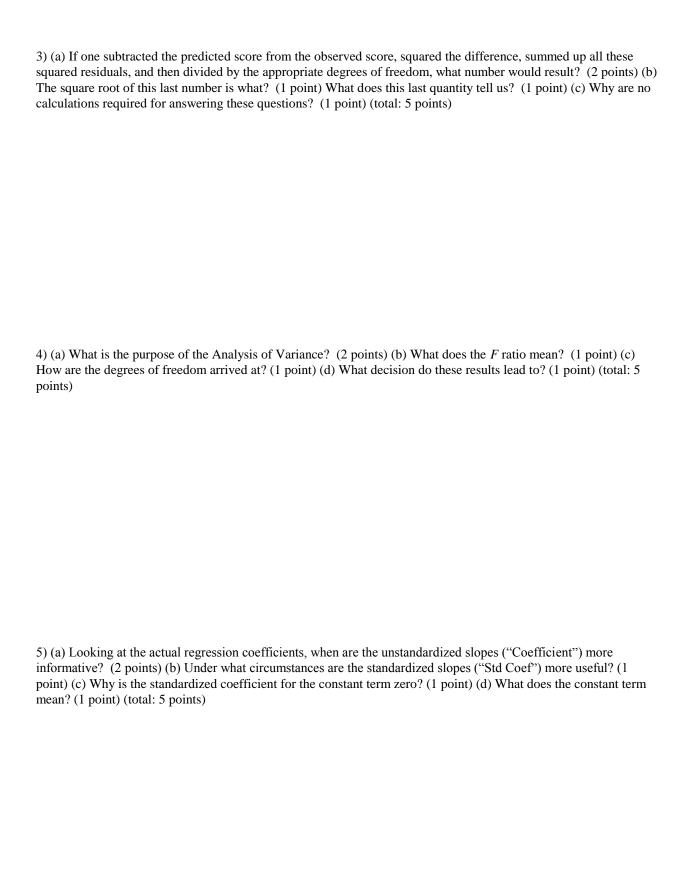
Psychology 204b (Causal Modeling) Simonton Winter 2008

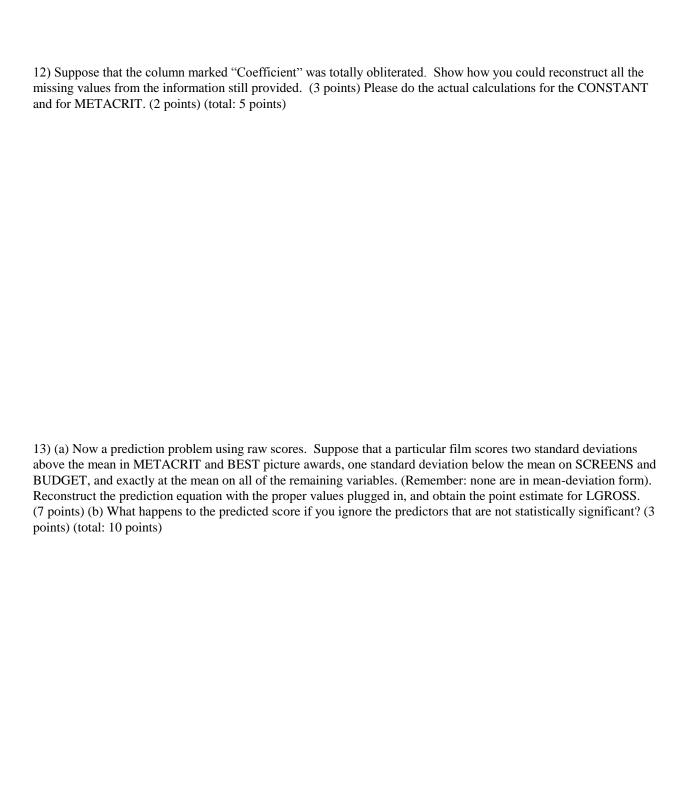
Exam I Your Name	
The statistics reported on the last page of this test are taken from a database regarding award-winning narrative film. The dependent variable LGROSS is a measure of the US box office earnings (in millions of dollars, log transformed). The independent variables are as follows: BUDGET = cost to make (in millions of dollars); SCREENS = the number of screens on which it opened; METACRIT = a composite of the critical reviews the film received on first theatrical release (on a 1-100 point scale); BEST = a measure of the number of best picture nominations and awards received from 7 major organizations; DRAMATIC, VISUAL, TECHNIC, and MUSIC similar tabulations for awards in the corresponding categories (e.g., DRAMATIC = directing, writing, acting, and editing). All variables are in raw-score form.	S
1) Use the results given in the printout to construct two separate regression equations, one for the original scores (3 points) and the other for the standardized scores (2 points). (total: 5 points)	

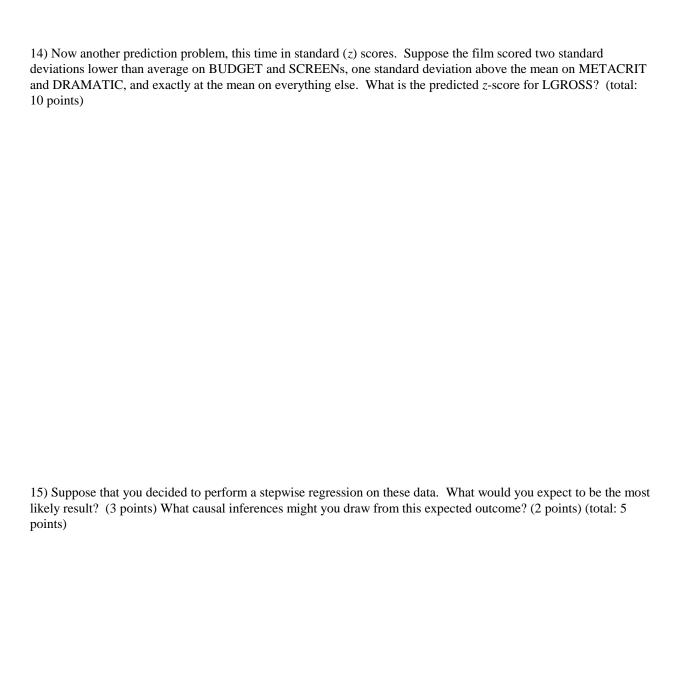
- 2) (a) What proportion of variance in the dependent variable is explained by the variables in the equation? (2 points)
- (b) Is there any difference between sample and population estimates in this respect? What does this mean? (1 point)
- (c) What is the correlation between the predicted value and the actual value? (2 points) (total: 5 points)

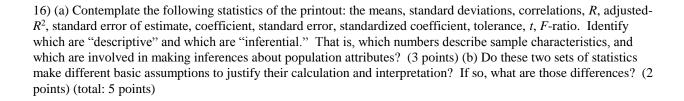


6) (a) What do the values in the column marked "Std Error" tell us? (2 points) (b) If these values were not given, how might they be calculated from other columns on the printout? (3 points) (total: 5 points)
7) (a) If the significance levels for the predictors were omitted from the printout, how could these be determined from the values under column "t"? (2 points) (b) If we had no table of t values, but instead had a table of Fs, what could we do then? (1 point) (c) If columns marked "t" and "P(2 Tail)" were both missing, how could we still determine that a given regression coefficient was statistically significant at the .05 level according to a 2-tailed test? (2 points) (total: 5 points)
8) (a) Using the available information, calculate the semi-partial (or part) correlation between LGROSS and BUDGET removing from the latter variable all the variance that it shares with the other independent variables in the equation. (2 points) (b) What does the square of this semi-partial tell you? (1 point) (c) How is the size of this semi-partial related to that of the corresponding standardized coefficient? Which coefficient is smaller than the other? Under what situations would these be the same? (1 point) (d) Which of these must be statistically significant given the results on the printout? (1 point) (total: 5 points)

9) (a) Now calculate the squared semi-partial correlations for all the independent variables in the equation. Add them all up. (2 points) (b) How does this sum compare with the squared multiple correlation? Smaller or the same? (1 point) (c) How does this sum compare with the sum of the squared zero-order correlations between the dependent variable and each independent variable? (1 point) (d) What do these comparisons signify? (1 point) (5 points)
10) (a) What information on the printout enables you to discern the presence of suppression? (3 points) (b) Are there any suppression effects? If so, provide a substantive interpretation. (2 points) (total: 5 points) 11) (a) What information on the printout enables you to determine whether multicollinearity is a major problem? (2
points) (b) What does the column labeled "Tolerance" tell us? (2 points) (c) What is the reciprocal of a variable's tolerance sometimes called? Why? (1 point) (total: 5 points)







17) (a) Suppose that you do not like this equation and want to estimate your own. Specifically, you want to predict LGROSS using just two independent variables, BUDGET and SCREENS (i.e., the only things the distributor knows prior to opening weekend). In your enthusiasm for hand-done regression, you calculate the following: the standardized and unstandardized regression coefficients, the constant term, and the squared multiple correlation. (8 points) (b) Is this equation better or worse than the one given? Any surprises? If so, what do they mean? Does this change your answer to #15? (2 points) (total: 10 points)

N of cases Minimum Maximum Mean Standard Dev	LGROSS 329 -0.6931 6.2136 3.5563 1.4629	BUDGET 329 0.0000 200.0000 49.6474 40.1095	SCREENS 329 0.0000 4163.0000 1764.7903 1351.6160	METACRIT 329 9.0000 98.0000 59.1611 20.3127	BEST 329 0.0000 1.8750 0.1292 0.3004
N of cases Minimum Maximum Mean Standard Dev	DRAMATIC 329 0.0000 7.5833 0.6859 1.3304	VISUAL 329 0.0000 5.2500 0.3505 0.8747	TECHNIC 329 0.0000 6.0000 0.3921 0.8363	MUSIC 329 0.0000 3.1250 0.1755 0.4032	
Pearson correlati	ion matrix				
LGROSS BUDGET SCREENS	LGROSS 1.0000 0.6269 0.6294	BUDGET 1.0000 0.6925	SCREENS	METACRIT	BEST
METACRIT BEST DRAMATIC VISUAL TECHNIC MUSIC	0.0933 0.1396 0.1502 0.2573 0.4215 0.3040	0.0323 -0.1341 -0.0446 -0.1010 0.1949 0.4330 0.2002	-0.3815 -0.2646 -0.3443 -0.0429 0.1996 0.0494	1.0000 0.4686 0.5016 0.2726 0.2578 0.3301	1.0000 0.7900 0.5772 0.4045 0.5362
DRAMATIC VISUAL TECHNIC	DRAMATIC 1.0000 0.4970 0.3525	VISUAL 1.0000 0.5660	TECHNIC	MUSIC	
MUSIC	0.4206	0.5818	0.4541	1.0000	

Number of observations: 329

Dep Var: LGROSS N: 329 Multiple R: 0.7912 Squared multiple R: 0.6259

Adjusted squared multiple R: 0.6166 Standard error of estimate: 0.9058

Effect	Coefficient S	td Error	Std Coef	Tolerance	t I	?(2 Tail)			
CONSTANT	0.7475	0.2158	0.0000		3.4641	0.0006			
BUDGET SCREENS	0.0075 0.0007	0.0019	0.2060 0.6477	0.4238	3.9224 11.9787	0.0001			
METACRIT	0.0162	0.0031	0.2244	0.6363	5.2344	0.0000			
BEST DRAMATIC	-0.3337 0.3183	0.3015	-0.0685 0.2895	0.3049	-1.1066 4.9169	0.2693			
VISUAL	0.0710	0.0832	0.0424	0.4719	0.8526	0.3945			
TECHNIC MUSIC	0.0562 0.1175	0.0829	0.0321	0.5210 0.5543	0.6783	0.4981			
Analysis of Variance									
Source	Sum-of-Squar	es df	Mean-Square	F-ratio	P				
Regression Residual	439.393 262.578		54.9241 0.8206	66.9351	0.0	0000			