Name Key
(Please ALSO write your name on the back of the last page.)

Ag 400 - Quiz 4
Makeup - Fall 2006

A. Suppose you are interested in ascertaining whether the rate of government taxation on the return (interest dividends, capital gains, etc.) from savings is related to people's willingness to save. The following data were compiled from eight different nations to address this question.

Country	(Y) Personal Savings Rate	(X) Taxation Rate
Italy	20	25
Japan	20	5
France	15	15
W. Germany	16	15
United Kingdom	13	20
Canada	10	32
Sweden	5	30
United States	5	50
Mean score	13	24
Sum of the scores	104	192
Sum of the squares of the scores	1600	5924
Sum of the squares of the deviations about the mean	248	1316
Sum of the cross-products (Σ_{xy})	-451	

(20) 1. Test the statistical significance of the linear relationship between personal savings rate and taxation rate. Use the .05 level and a two-tailed test to determine statistical significance. Specify the following:

a. State the Null Hypothesis and the Alternative Hypothesis.

H_0: there is no linear relationship between Personal Savings Rate and Taxation Rate

H_A: there is a linear rel...

b. Show the necessary calculations.

$$r = \frac{-451}{\sqrt{(248)(1316)}} = -.789$$

c. Specify degrees of freedom.

df = 6

d. Report the critical value of r.

e. Reject or do not reject the Null Hypothesis.

Reject H_0

f. Interpret the meaning of your conclusion in (e) precisely in terms of the problem.

There is a negative linear relationship between Personal Savings Rate + Taxation Rate

(10) 2. Estimate the Taxation Rate for a country where the personal savings rate is 12. Show your work.

$X = a_{xy} + b_{xy} Y$ $\quad b_{xy} = \frac{\Sigma xy}{\Sigma y^2} = \frac{-451}{248} = -1.81855$

$a_{xy} = \bar{X} - b_{xy}\bar{Y} = 24 + 23.641127 = 47.64$

$X = -1.8185(12) + 47.64$

$X = 47.64 - 1.81855(12) = 25.82$

(10). 3. Estimate the personal Savings rate for a country where the Taxation Rate is 23.

$Y = a_{yx} + b_{yx} X$ $\quad b_{yx} = \frac{-451}{1316} = -.3427$

$a_{yx} = \bar{Y} - b_{yx}\bar{X} = 13 + (.3427)(24) = 21.2248$

$Y = 21.2248 - .3427(23) = 13.3427$

(10) 4. Answer the following by indicating whether each statement is true (T) or false (F) in terms of the above problem. If any part of a statement is untrue, it should be marked false (F). Add comments, if you wish, to clarify your answers.

T a. The statistical unit here is a country.

T b. Savings Rate and Taxation Rate are treated as interval scales in the above analysis.

F c. Based on the above analysis, for every 1 point increase in taxation rate, you should predict b_{yx} points change in ~~Taxation Rates~~ Personal Savings Rate.

T d. Using the above model, if a country's Taxation Rate was equal to 24, you should estimate its Personal Savings Rate to be 13; if the country's Personal Savings Rate was 13, you should estimate its Taxation Rate to be 24.

T e. r^2_{yx} measures the proportion of the variation in Taxation Rates in the sample that is accounted for by variation in Personal Savings Rates.

B. Suppose that you are interested in assessing the relationships of age and education to the social conservatism of people living in Mudville. You survey a random sample of Mudville residents and obtain information on the following variables.

Conservatism Score (score)
This is an attitude scale developed to assess the level of social conservatism of individuals. Possible scores range from zero to 100. The higher the score, the more conservative the individual's social views.

Age in years of respondent (ageyrs)

Years of formal schooling completed by respondent (ed)

(50) The attached output was obtained using SPSS and the data set described above. Indicate whether each of the following statements is true (T) or false (F) in terms of the output. If any part of a statement is untrue, it should be marked false (F). Add comments, if you wish, to clarify your answers.

T 1. The data file consists of 375 cases of which two cases do not have information on the respondent's conservatism score. These two cases are deleted from all of the statistical analysis on the output.

T 2. The statistical unit is a Mudville resident; and there are 375 lines in the SPSS data file used in this analysis.

T 3. In this sample, the linear correlation between age of the respondent and his/her conservatism scores is positive. Overall, as age increases, conservatism scores increase.

F 4. Based on this output, if you reject the following Null Hypothesis in favor of the following Alternative Hypothesis, $\alpha < .001$, $\beta = 0.00$

H_o: In Mudville population there is no significant linear relationship between age and Conservatism Score.

H_A: There is a significant linear relationship between age and Conservatism Score in Mudville.

T 5. The following regression equation accounts for 8.6% of the variation in Conservatism Scores in this sample.

$$\text{Score} = 5.0 + 1.446\ (\text{ageyrs}) - .010\ (\text{ageyrs})^2$$

T 6. The curvilinear (quadratic) correlation coefficient for the relationship of age in years to Conservatism Score is larger than is the correlation coefficient for the linear relationship of age in years to Conservatism Score. Linear correlation coefficients are <u>never</u> greater than the (quadratic) curvilinear correlation coefficients using the same data.

F 7. Based on this analysis, we can conclude that the quadratic curve is a significantly ($p<.05$) better fit to the data than is a straight line for describing and predicting Conservatism Scores from age in years.

T 8. Overall, in this sample, older respondents have fewer years of formal schooling than do younger respondents.

F 9. The zero-order correlation between number of years of formal schooling and Conservatism Scores is negative and statistically significant at the .001 level.

T 10. In the bivariate analysis, years of formal schooling accounts for about 2.2% of the variation in Conservatism Scores in the sample.

T 11. The interrelationship between "ed" and "ageyrs" is statistically significant at the .001 level using a 2-tailed test.

F 12. Overall, for every one additional year of formal schooling, Conservatism Scores would be predicted to decline by .149 points.

F 13. The Multiple R reported on Page 6 is positive in this analysis because the positive correlation of "ageyrs" to "score" is greater than the negative correlation of "ed" to "score".

F 14. Approximately 29.2% of the total variation in Conservatism Scores is accounted for by the combined linear effects of age in years and years of formal education.

F 15. For these data, the Multiple Coefficient of Nondetermination equals .708. $1-R^2$

T 16. Adjusted R Square on Page 6 of the output is never negative and never greater than R Square for the same data.

F 17. R and R Square on Page 6 are statistically significant at the .001 level. Thus, we can conclude that both age and years of formal schooling are significant linear correlates of Conservatism Score.

T 18. The multiple regression equation relating age in years and years of formal schooling to Conservatism Score obtained from this analysis is as follows:

score = 39.852 + .508 (ageyrs) - 1.120 (ed)

F 19. On Page 6 the "B" for ed (-1.120) is greater in absolute value then is the "B" value for ageyrs (+.508). Thus, you can conclude that years of education has a stronger relationship to Conservatism Scores than does age in years.

T 20. Adjusting for years of formal schooling and using the results of the analysis on Page 6, you should predict that a 40-year old man would have a Conservatism score that was about 3.0 points higher than that of a 46-year old man.

T 21. If "ageyrs" were deleted from the regression model on Page 6, the proportion of the variation in Conservatism Scores in the sample accounted for by the model would decline from .085 to .022. This decline is statistically significant at the .001 level.

F 22. From this output, you can conclude that 25.5% of the variation in Conservatism Scores that is not associated with variation in years of formal schooling is accounted for by age of the respondent.

T 23. In the bivariate analysis, the effect of "ed" was inflated by a failure to control for the effect of "ageyrs." When "ageyrs" was controlled, the strength of the relationship of "ed" to "score" declined to nonsignificance ($p>.05$).

F 24. In this analysis the strengths of the partial relationships of "ed" to "score" and "ageyrs" to "score" were both less their corresponding zero order relationships to "score." Partial correlations are never greater than their corresponding zero order correlations.

F 25. Based on this analysis, the parsimonious model for predicting Conservatism Scores from age and years of formal schooling is:

Score = 39.852 + .508 (ageyrs)

Curve Fit

Model Description

Model Name		MOD_13
Dependent Variable	1	score
Equation	1	Linear
	2	Quadratic
Independent Variable		ageyrs
Constant		Included
Variable Whose Values Label Observations in Plots		Unspecified
Tolerance for Entering Terms in Equations		.0001

Case Processing Summary

	N
Total Cases	375
Excluded Cases[a]	2
Forecasted Cases	0
Newly Created Cases	0

a. Cases with a missing value in any variable are excluded from the analysis.

Variable Processing Summary

		Variables	
		Dependent	Independent
		score	ageyrs
Number of Positive Values		373	375
Number of Zeros		0	0
Number of Negative Values		0	0
Number of Missing Values	User-Missing	0	0
	System-Missing	2	0

score

Linear

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
.283	.080	.078	18.111

The independent variable is ageyrs.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	10595.545	1	10595.545	32.301	.000
Residual	121696.89	371	328.024		
Total	132292.43	372			

The independent variable is ageyrs.

Coefficients

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
ageyrs	.549	.097	.283	5.683	.000
(Constant)	24.123	4.663		5.174	.000

Quadratic

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
.293	.086	.081	18.080

The independent variable is ageyrs.

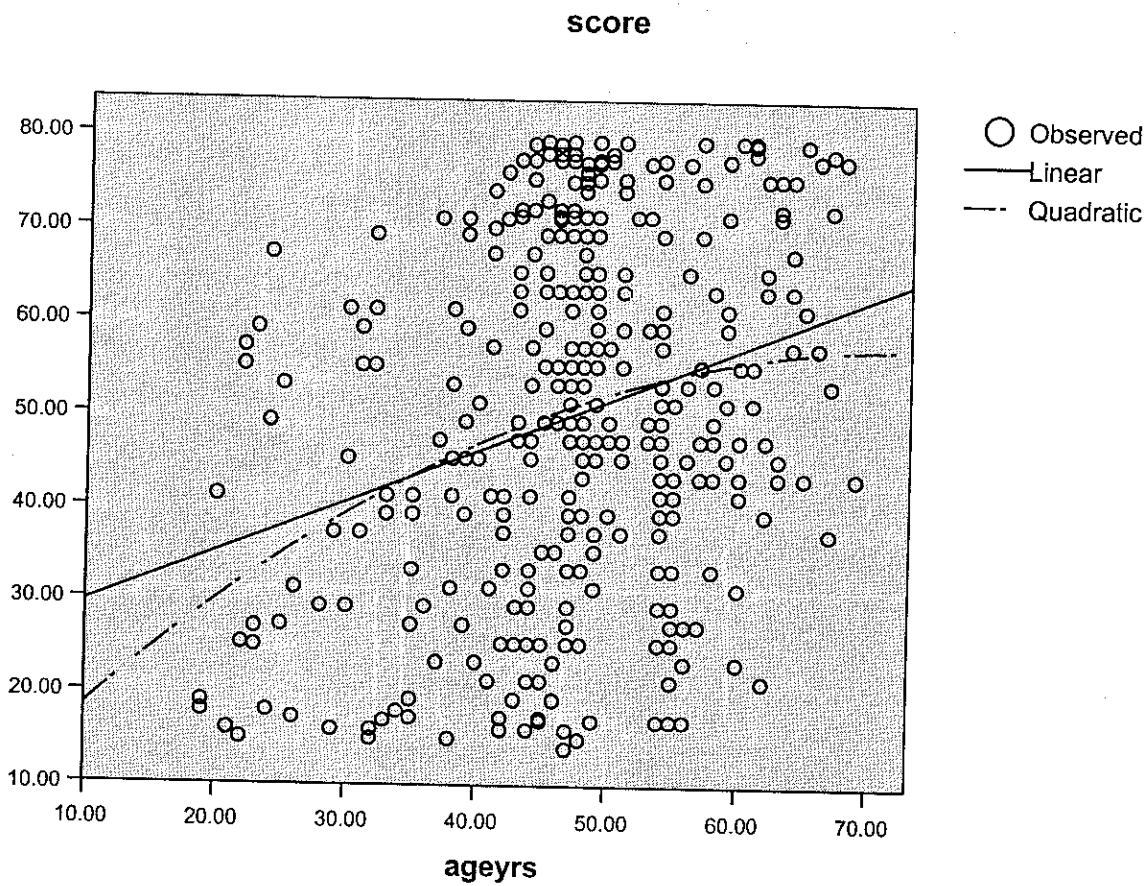
ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	11350.553	2	5675.276	17.362	.000
Residual	120941.88	370	326.870		
Total	132292.43	372			

The independent variable is ageyrs.

Coefficients

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
ageyrs	1.446	.599	.746	2.416	.016
ageyrs ** 2	-.010	.007	-.469	-1.520	.129
(Constant)	5.000	13.416		.373	.710


score
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
ageyrs
Observed
Linear
Quadratic

```
REGRESSION
  /DESCRIPTIVES MEAN STDDEV CORR SIG N
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA ZPP
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT score
  /METHOD=ENTER ageyrs ed  .
```

Regression

Descriptive Statistics

	Mean	Std. Deviation	N
score	50.0811	18.85802	373
ageyrs	47.3217	9.72932	373
ed	12.3123	1.28668	373

Correlations

		score	ageyrs	ed
Pearson Correlation	score	1.000	.283	-.149
	ageyrs	.283	1.000	-.276
	ed	-.149	-.276	1.000
Sig. (1-tailed)	score	.	.000	.002
	ageyrs	.000	.	.000
	ed	.002	.000	.
N	score	373	373	373
	ageyrs	373	373	373
	ed	373	373	373

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	ed, ageyrs[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: score

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.292[a]	.085	.081	18.08259

a. Predictors: (Constant), ed, ageyrs

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	11309.850	2	5654.925	17.294	.000[a]
	Residual	120982.58	370	326.980		
	Total	132292.43	372			

a. Predictors: (Constant), ed, ageyrs

b. Dependent Variable: score

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	39.852	11.616		3.431	.001			
	ageyrs	.508	.100	.262	5.064	.000	.283	.255	.252
	ed	-1.120	.758	-.076	-1.478	.140	-.149	-.077	-.073

a. Dependent Variable: score