Statistics Midterm 9803
ANOVA, correlation, and multiple regression, and *t* tests

Name: Benedict Tieniber

"Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write." H. G. Wells.

PART I (5 points)

<u>True and False – Add short explanation to your selection</u>

1. **One of the main objectives of simple linear regression is to determine the type of relationship that exists between X and Y and test for its significance**
   TRUE

2. **If you want to know if your salary depend on your college GPA. Then, College GPA in this case is the dependent variable.**
   FALSE – The dependent variable is salary

3. **We want to test whether ex husbands and ex wives have the same reasons to justify the divorce? This is pair sample t-test**
   FALSE – Independent T-Test

4. **In the article "Students Perceptions of satisfaction and anxiety on online doctoral program…" The authors concluded (p. 91) "Students with lower technological anxiety scores experienced higher level of satisfaction" – They performed a one-tailed test correlation.**
   FALSE – The test used was the two-tailed test

5. **In the article "Do perceptions of college students from one liberal arts college on long island vary…" The authors show how wealthy students outperformed poor in the uses of spreadsheets.**
   FALSE – Wealthy was never used in the study

Section 2 (5 points)

1. **Research about effect size and give an example about it. Why is this important? Effect size = r^2 for ANOVA. The 'specific' percent of variance accounts for the change.**

2. **What's the difference between p-value and effect size?**
   The difference between p-value and effect size is p-value is significance and effect size is the percent of variance the independent variable has on the dependent variable.

3. **A researcher used correlation to examine the relationships between a binary predictor variables (e.g. Asian American versus Caucasian) and a continuous criterion variable (AP physics score).**
   He coded:
   Asian American (AA) = 1
   Caucasian (C) = 2

   Please explain the following results:

   For ethnic (AA=1, C=2) and AP physics score: r = .34, p = .054
   No correlation/relationship between Caucasians and Asian Americans.

   For ethnic (AA=1, C=2) and AP physics score: r = -0.6 , p = .001
   The correlation is negative. Asian Americans are performing better in Physics than Caucasians.

4. **<u>Matching (2 points)</u>**

- **Paired Sample T-Test** **-Involves pretest and post-test**
- **ANOVA** **-Involves Mean differences**
- **Two-way ANOVA** **-Involves interaction among 2 variables or factors**
- **Binary Linear Regression** **Addresses relationship between 2 variables**

Part III (10 points)

2. Multiregression and correlation (5 points)

Based on the data set from the Multiple regression chapter , We conducted a multiregression analysis to predict the overall injury index from previous medical difficulties and age

Dependent Variable = Injury
Independent Variables (7 variables) = medimex, age, abdoms, gluts, quads, grip, arms

- **We checked distributions → all the variables showed normal distribution**
- **Table 1 shows the bivariate correlations between independent variables and dependent variable**
- **Table 2 shows the multiregression results**

		INJURY	QUAD
INJURY	Pearson Correlation	1	-.1
	Sig. (2-tailed)	.	.1
	N	100	1
QUADS	Pearson Correlation	-.162	
	Sig. (2-tailed)	.107	
	N	100	1
GLUTS	Pearson Correlation	-.393**	.4
	Sig. (2-tailed)	.000	.0
	N	100	1
ABDOMS	Pearson Correlation	-.232*	.2
	Sig. (2-tailed)	.020	.0
	N	100	1
ARMS	Pearson Correlation	-.243*	.3
	Sig. (2-tailed)	.015	.0
	N	100	1
GRIP	Pearson Correlation	-.099	.1
	Sig. (2-tailed)	.328	.0
	N	100	1
AGE	Pearson Correlation	.290**	.0
	Sig. (2-tailed)	.003	.3
	N	100	1
MEDINDEX	Pearson Correlation	.337**	-.1
	Sig. (2-tailed)	.001	.1
	N	100	1

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

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

Table 1. Bivariate Correlations (also call zero-order correlations)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity S	
		B	Std. Error	Beta			Tolerance	
1	(Constant)	-141.553	101.652		-1.393	.167		
	QUADS	.540	.600	.100	.900	.371	.609	
	GLUTS	-3.386	1.042	-.375	-3.250	.002	.562	
	ABDOMS	-.682	.627	-.117	-1.087	.280	.646	
	ARMS	-.842	.657	-.138	-1.281	.203	.649	
	GRIP	.805	1.005	.081	.801	.425	.741	
	AGE	5.331	1.501	.330	3.553	.001	.867	
	MEDINDEX	89.909	99.132	.090	.907	.367	.755	

a. Dependent Variable: INJURY

Questions:

1) From table 2,

a) write the multiregression equation based on the unstandardized coefficients .54 quads – 3.386 gluts - .682 abdomen - .842 arms + .805 grip + 5.331 age – 89.909 medium index – 141.553 = Injury

INJURY =

b) Interpretation of betas:

- **Interpret "b" for quads -- for each 1-unit increase in quads, injury is expected to up by .54 when holding all other variables constant.**
- **Interpret "b" for gluts -- for each 1-unit increase in support, injury is expected to down by 3.386, when holding all other variables constant**
- **Interpret "constant" -- if a person has a score of "0" on all predictors, their injury is expected to be -141.553.**
- **The 2 best predictors of the injury index are (Tip, check magnitude of betas, and p-values): Gluts and age**
- **The 2 worst predictors of the injury index are: Grip and medical index**

2) **There are 5 patterns of bivariate/multiregression relationship as shown in Table 3:**

There are 5 patterns of bivariate/multivariate relationship

Multiple regression weight	Simple correlation with the criterion: -	Simple correlation with the criterion: 0	Simple correlation with the criterion: +
-	Bivariate relationship and multivariate contribution (to this model) have same sign	"Suppressor variable" – no bivariate relationship but contributes (to this model)	"Suppressor variable" – bivariate relationship & multivariate contribution (to this model) have different signs
0	Non-contributing – probably because colinearity with one or more other predictors	Non-contributing – probably because of weak relationship with the criterion	Non-contributing – probably because colinearity with one or more other predictors
+	"Suppressor variable" – bivariate relationship & multivariate contribution (to this model) have different signs	"Suppressor variable" – no bivariate relationship but contributes (to this model)	Bivariate relationship and multivariate contribution (to this model) have same sign

c) **Explain the relationship between each independent variables and the model (Tip: You need the information on Tables 1, 2 and 3).**
Quads 0,0 --- non-contributing
Gluts -,- --- bivariate
Abdominals -,0 --- non-contributing
Arms --, 0 --- non-contributing
Grip 0,0 --- non-contributing
Age +,+ ---- bivariate
Medical Index +,0 --- non-contributing

d) **Which variable(s) are/is causing collinearity? (if any) Explain collinearity.**
Arms, abs, and medical index. Variables have an effect – one 'eats' the other.