SAMPLE QUESTIONS – STAT 2

COMPREHENSIVE EXAM SAMPLE QUESTIONS – ONEWAY ANOVA

Does coffee help people become sober more quickly after drinking too much? A sample of 40 volunteers were randomly assigned to one of four groups of 10 subjects each. One of these groups serves as a control and receives no alcohol. Subjects in each of the other three groups drink a fixed amount of alcohol in a 1-hour period. During the next half-hour, subjects in the second group drink two cups of decaffeinated coffee, subjects in the third group drink two cups of regular coffee, and subjects in the fourth group drink two cups of water. Finally, all subjects are given a reaction-time test to determine mental alertness with the lower the score on the test, the quicker the reaction time and thus greater alertness. The attached printout gives the results of the analysis of the data. Use $\alpha = .05$ to determine statistical significance.

(a) What are the independent and dependent variables for this study?

(b) Why did the researcher randomly assign the subjects to the four groups?

(c) Prior to examining whether the group means differ it is necessary to test the assumption of homogeneity of variance.
(1) Do we meet this assumption?
(2) What specific information on the printout did you use to come to this conclusion?
(3) What are the other assumptions that underlie the application of analysis of variance procedures?

(d) (1) Is the average reaction time the same for all four groups?
(2) What specific information on the printout did you use to come to your conclusion?

(e) The $\omega^2 = .186$ for these data. What does this represent?

(f) According to the Tukey results on the attached printout, what pairs of groups differ?

(g) To whom, if anyone, can we generalize our findings?

(h) Write a brief Results section describing the findings of this experiment based on the analyses you have just completed. You do NOT have to present results in tables.

```
ONEWAY
  react BY sober
  /STATISTICS DESCRIPTIVES HOMOGENEITY
  /MISSING ANALYSIS
  /POSTHOC = TUKEY ALPHA(.05).
```

Oneway

Descriptives

react

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
					Lower Bound	Upper Bound
1.00 control	10	194.9000	10.80586	3.41711	187.1700	202.6300
2.00 decaf	10	212.3000	12.89315	4.07717	203.0768	221.5232
3.00 regular	10	211.1000	14.13781	4.47077	200.9864	221.2136
4.00 water	10	211.5000	14.56976	4.60736	201.0774	221.9226
Total	40	207.4500	14.64442	2.31549	202.7665	212.1335

Descriptives

react

	Minimum	Maximum
1.00 control	179.00	215.00
2.00 decaf	191.00	230.00
3.00 regular	189.00	231.00
4.00 water	192.00	231.00
Total	179.00	231.00

Test of Homogeneity of Variances

react

Levene Statistic	df1	df2	Sig.
.715	3	36	.550

ANOVA

react

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2107.500	3	702.500	4.042	.014
Within Groups	6256.400	36	173.789		
Total	8363.900	39			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: react

Tukey HSD

(I) sober	(J) sober	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00 control	2.00 decaf	-17.40000*	5.89557	.027	-33.2781	-1.5219
	3.00 regular	-16.20000*	5.89557	.044	-32.0781	-.3219
	4.00 water	-16.60000*	5.89557	.038	-32.4781	-.7219
2.00 decaf	1.00 control	17.40000*	5.89557	.027	1.5219	33.2781
	3.00 regular	1.20000	5.89557	.997	-14.6781	17.0781
	4.00 water	.80000	5.89557	.999	-15.0781	16.6781
3.00 regular	1.00 control	16.20000*	5.89557	.044	.3219	32.0781
	2.00 decaf	-1.20000	5.89557	.997	-17.0781	14.6781
	4.00 water	-.40000	5.89557	1.000	-16.2781	15.4781
4.00 water	1.00 control	16.60000*	5.89557	.038	.7219	32.4781
	2.00 decaf	-.80000	5.89557	.999	-16.6781	15.0781
	3.00 regular	.40000	5.89557	1.000	-15.4781	16.2781

*. The mean difference is significant at the .05 level.

Homogeneous Subsets

react

Tukey HSD[a]

sober	N	Subset for alpha = .05	
		1	2
1.00 control	10	194.9000	
3.00 regular	10		211.1000
4.00 water	10		211.5000
2.00 decaf	10		212.3000
Sig.		1.000	.997

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 10.000.

SAMPLE COMPREHENSIVE EXAM QUESTIONS – MULTIPLE REGRESSION

A researcher in the area of child development is interested in the relative influences on student achievement of family background, family socialization practices, students' achievement values, and students' self-concepts. She gathered data on 200 junior high school students and that data was analyzed using multiple regression. The attached printout shows the results of her analysis. In that analysis, she regressed student achievement (**ACHIEVE**) on family background (**SES**), ethnicity (**ETHNIC** coded 0 = White, 1 = Hispanic), two measures of family socialization practices (independence training - **INDEPEND** - and family emphasis on achievement - **EMPHASIS**), student self-concept (**SELFCON**) and student achievement values (**VALUES**). Use the output and $\alpha = .05$ to answer the following questions:

1. Does multicollinearity appear to be a problem in the analyses? Why or why not?

2. Does the set of independent variables explain a significant proportion of variance in student achievement? What information on the printout did you use to answer this question.

3. What proportion of variance in student achievement is explained by the set of independent variables? What did you use to determine your answer?

4. Which of the independent variables have a significant influence on student achievement? What particular information on the printout did you use to answer this question?

5. What is the relative importance of the independent variables in their influence on student achievement? What particular information on the printout did you use to answer this question?

6. Are family socialization practices (EMPHASIS and INDEPEND) more important than family background (SES) in influencing student achievement? Why or why not?

7. How do you **substantively** interpret the coefficient for ETHNIC? Be specific.

Regression

Descriptive Statistics

	Mean	Std. Deviation	N
sex	.514000	.5090000	200
ethnic	.326000	.4720000	200
ses	2.926000	2.3530000	200
independ	6.218000	2.5520000	200
emphasis	-.011000	.9510000	200
values	-.021000	.8460000	200
selfcon	-.023000	.8800000	200
achieve	45.524000	11.7850000	200

Correlations

		sex	ethnic	ses	independ	emphasis	values	selfcon	achieve
Pearson Correlation	sex	1.000	.057	-.141	.260	-.148	.151	-.076	.011
	ethnic	.057	1.000	.652	.234	.017	.436	.302	.473
	ses	-.141	.652	1.000	.264	.170	.245	.278	.441
	independ	.260	.234	.264	1.000	.166	.371	.145	.376
	emphasis	-.148	.017	.170	.166	1.000	-.243	.106	-.092
	values	.151	.436	.245	.371	-.243	1.000	.294	.381
	selfcon	-.076	.302	.278	.145	.106	.294	1.000	.472
	achieve	.011	.473	.441	.376	-.092	.381	.472	1.000

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.674[a]	.454	.437	8.8393603

a. Predictors: (Constant), selfcon, emphasis, independ, ethnic, values, ses

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	12558.441	6	2093.073	26.788	.000[a]
	Residual	15079.918	193	78.134		
	Total	27638.359	199			

a. Predictors: (Constant), selfcon, emphasis, independ, ethnic, values, ses

b. Dependent Variable: achieve

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	33.702	1.938		17.392	.000		
	ethnic	4.643	1.898	.186	2.446	.015	.489	2.044
	ses	.940	.364	.188	2.583	.011	.535	1.869
	independ	1.228	.279	.266	4.407	.000	.776	1.288
	emphasis	-2.567	.732	-.207	-3.508	.001	.811	1.233
	values	.044	.950	.003	.046	.963	.608	1.644
	selfcon	4.635	.776	.346	5.976	.000	.843	1.187

a. Dependent Variable: achieve

SAMPLE RESEARCH SCENARIOS

For each of the following scenarios, determine the most appropriate statistical procedure that would be used to answer the research question:

An experimental psychologist wants to test the hypothesis that memory for pictures is better than memory for words. The psychologist performs an experiment in which one group of students view 30 slides with words and another group views 30 slides with pictures. Students are then given a recall test.

An educational researcher is interested in determining the relative influences of socioeconomic background, educational aspirations, ability, and gender on academic achievement.

The federal government is interested in testing whether an advertising campaign for gasoline conservation is effective. For a sample of subjects, they record the amount of gasoline used in a one month period prior to the advertising campaign and then for one month following the campaign.

A nutritionist wants to determine whether there are differences in the sugar content of three different breakfast cereals.

A sales manager is interested in whether there is a relationship between the amount of money he spends on advertising each month and his net profit each month. He gathers data each month for two years.